KUHNS, JACK, The Role of Schematic Support in Strategy Choice during Cognitive Skill Learning. (2018)
Directed by Dr. Dayna R. Touron. 104 pp.

This study examines influences on strategic differences in skill learning that occur with increasing age. Older adults differ in their strategic approach to cognitive skill acquisition tasks, where their progression from slow algorithmic processing to faster memory-based processing is slowed relative to young adults. In addition to difficulties older adults have with learning new associations, the difference in task approach has been linked to strategic choice, where factors such as lower confidence change how they interact with the task (e.g., Touron, 2015). The present study sought to understand older adults' strategic task approach by manipulating the task items to be more naturalistic with everyday experience. Participants completed a task that associated grocery items with prices, which are easier to learn if the prices are consistent with everyday experience (Castel, 2005). The relations between the grocery items and prices were manipulated to be familiar by approximating market prices, or to be unfamiliar by being overpriced. I found that use of the market-prices facilitates older adults' strategic approach to the task, demonstrated through greater and earlier use of memory-based processing than older adults with overpriced items. Surprisingly, the young adults in the overpriced condition also showed less use of memory-based processing, linked with lower task confidence; young adults have not previously shown reluctance to use memory strategies in cognitive skill acquisition tasks. Consequences of task confidence are discussed, as well as implications for theories of cognitive skill acquisition.

Keywords: aging, automaticity, skill acquisition, strategies, metacognition

THE ROLE OF SCHEMATIC SUPPORT IN STRATEGY CHOICE DURING

COGNITIVE SKILL LEARNING

by

Jack Kuhns

A Thesis Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Greensboro
2018

Approved by

____
Committee Chair

APPROVAL PAGE

This thesis written by JACK KUHNS has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair ______________________________

Committee Members ______________________________

____

____

____

____
Date of Acceptance by Committee

____
Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Old age is a time fraught with many changes that are often viewed as debilitating, inevitable, and out of one's control. Besides concerns of declining health and mobility, changes in memory are important for many older adults. Learning new associations is particularly impaired with increasing age (e.g., Kausler, 1994; Naveh-Benjamin, 2000), and negative beliefs about one's memory ability may also decrease the efficiency of learning new associations (Touron & Hertzog, 2004a). I will first review age-differences in memory performance, which are often attributed as the cause for differences in strategic behavior in learning tasks, then review age-differences in strategic behavior in simple memory tasks and finally age-differences in cognitive skill acquisition tasks.

Age-related deficits in associative memory have been well-documented and are a subject of great interest in cognitive aging research (e.g., Kausler, 1994). Chalfonte and Johnson (1996) suggested that this deficit may in part be due to older adults' difficulty binding information into complex memories. Naveh-Benjamin (2000) extended and tested this idea by comparing memory for associations to memory for single items. Older adults' memory for single words and related word pairs showed no deficits when compared with the young, but their cued recall for non-words and unrelated word pairs was particularly impaired, reaching 25-50% of young adults' recall. These patterns of results supported Naveh-Benjamin's (2000) Associative Deficit Hypothesis (ADH), that older adults' decreased episodic memory is due in part to decreased ability to encode and retrieve associations among units of information or attributes within events.

When older adults' show preservation of memories for words and related items, Naveh-Benjamin (2000) credits this to being able to rely on preexisting knowledge to reduce processing costs at encoding. Effective remembering for arbitrary information requires a more deliberate and meaningful encoding strategy, which older adults are less likely to spontaneously adopt, i.e., without explicit instruction (e.g., Hertzog & Hultsch, 2000). Often older adults instead rely on incidental learning, which is unfortunately another impaired memory domain for older adults (Schneider & Pressley, 1997), potentially due to reduced processing efficiency or speed (Salthouse, 1996).

It's been suggested that associative deficits are due to production deficiencies in implementing successful mediational strategies, where participants generate a mediator linking the two associative items (e.g., an interactive image of the items, or a sentence containing both items; Kuhlmann & Touron, 2012), which are more successful in aiding memory than simple rote repetition. Indeed, older adults are often less likely to implement successful mnemonic strategies than young adults, who often implement them without instruction (e.g., Kausler, 1994). However, production deficiencies are not likely the culprit behind age-related associative memory changes.

Simply instructing older adults to use effective mediational strategies can eliminate age-differences in memory-strategy use (Dunlosky & Hertzog, 1998, 2001; Naveh-Benjamin, Brav, & Levy, 2007), although robust age differences in recall are still found for paired-associates even when older adults are instructed to, and use, potentially effective strategies (Dunlosky & Hertzog, 1998). This speaks to a problem older adults have with utilizing and/or recalling the mediators used (Dunlosky, Hertzog, & Powell-Moman, 2005; Kuhlmann & Touron, 2012; Naveh-Benjamin et al., 2007).

Older adults instructed to recall mediators used at study in order to help their memory at test have been shown to outperform older adults not instructed to recall their mediators at test (Naveh-Benjamin et al., 2007), suggesting that older adults' retrieval strategies aren't always efficient (for example, not attempting to recall the mediator of a word-pair in one's memory search if the target does not immediately come to mind). Dunlosky et al. (2005) found that when mediators generated during study were correctly recalled at test, older adults performed equivalently with young adults, but when mediators were only partially recalled, older adults were more likely to forget the cued response. Although older adults have been shown to utilize and produce mnemonic strategies equivalently to young adults, they don't always benefit equivalently from doing so.

Metacognitive Influences on Task Performance

Even though differences have been found in associative memory for older adults, this doesn't necessarily mean that older adults are deficient in evaluating the contents of their memory (Hertzog & Hultsch, 2000). Much work has studied whether older adults' task-related monitoring of their memory (i.e., metacognitive monitoring) is intact, which it largely is (Hertzog & Hultsch, 2000, but see Hertzog, 2015 for a review of boundary conditions). However, age-related differences are broadly found in the domain of metacognitive control, or the use of metacognitive monitoring processes to achieve strategic control over cognition (e.g., Hertzog, 2015; Nelson & Narens, 1990).

In a series of early studies investigating age differences in metacognition, Murphy, Gabreisheski, Schmitt, and Sanders (1981) tested young and older adults in a recall-readiness task, where participants were asked to learn a list of words just longer than their short-term memory capacity, and were given as much time to do so before indicating that they were prepared to recall the entire list. Though not all people recalled all the items from the list, older adults recalled fewer items than young adults, despite indicating that they were ready to recall all the items. Further, older adults were much less likely to implement self-instantiated testing strategies to ensure that they knew the words (a form of retrieval practice, e.g., Pyc & Rawson, 2009).

In a follow-up study, Murphy, Schmitt, Caruso, and Sanders (1987) explicitly asked young and older adults to use this self-testing strategy (versus a control condition with no strategy instructions), and found that the strategy-instructed older adults performed just as well as younger adults, and out-performed older adults under no strategy instruction. This demonstrates a metacognitive control deficit, because older adults are just as capable of implementing the self-testing strategy when instructed, so those not instructed were using less effective encoding strategies. The act of self-testing lets one know what words they do and do not know, which can then be given special attention to ensure adequate encoding for the later test. Prior to signaling readiness for the recall test, one might ask themselves “Can I recall this list?” vs. “Can I recall each item on this list?”; the former question affords only familiarity of the list generally, while the latter sets one up to attempt recall for each item prior to the test. Older adults can perform as well as young adults in memory tests when given the time – as Murphy et al. (1987) show – but in these cases their control-based monitoring does not often yield effective study strategies.

Bottiroli, Dunlosky, Guerini, Cavallini and Hertzog (2010) replicated the basic finding of Murphy et al. (1987) where older adults self-tested significantly less than young adults, but called in to question the method used to assess metacognitive control, noting that the way tasks are presented to participants affords different levels of spontaneous strategy production. Murphy et al. presented their word lists affixed to a board, an unusual format for presenting word lists, which limits the explanation that age per se was the culprit for not producing the self-testing strategy. Bottiroli et al. (2010) explained that the board-presentation format limits one's strategy space, where older adults either did not even think of self-testing or did not figure out how to self-test in that format. To test this assumption, they presented single and paired associates on a board and on index cards. They did not find age differences in self-testing except for single words affixed to a board, with paired-associate index cards showing the highest level of self-testing overall across both age groups. These results support the important role that task features play in strategy production, and should factor importantly in studies of age-related strategy differences (cf. Kuhlmann & Touron, 2016; Touron & Hertzog, 2004b).

How a task is presented (including the kind of stimuli) can also have substantial influences on older adults' memory performance (e.g., Hess & Emery, 2012; Zacks, Hasher, & Li, 2000). Rahhal and Hasher (1998) presented instructions before a trivia game emphasizing that the goal was knowledge acquisition rather than just remembering the factoids for the later test, and found that standard age differences were alleviated, assuming the goal of knowledge acquisition motivated the older adults to study the material more deeply (which would enable longer-term retention) than just for a later test. Differences in story recall also support the notion that encoding approaches differ with age. Older adults tend to retell stories in an integrative, meaningful, and succinct style more so than the young, who tend to retell stories in exact detail (Adams, Smith, Nyquist, & Perlmutter, 1997). This work suggests that older adults might have different processing goals than young adults in memory tasks, resulting in qualitatively different memory traces than that of the young, which can be obscured in normal memory tasks looking for exact reproductions of the stimuli.

Hess and colleagues (see Hess & Emery, 2012 for a review of this work) have argued that the concordance between goal states that an older adult holds and task demands can largely influence older adults' memory performance through motivation and processing style. Hess' selective engagement hypothesis states that in situations of strong personal relevance and implication (e.g., in situations of social accountability where one is answering questions in front of other participants or confederates), standard age differences in memory recall would be attenuated, where the goal of appearing competent in front of others enhances their motivation for success. Indeed, this has been found in studies that manipulate motivation by means of social accountability, where participants must recall memoranda in front of others. In these studies, motivation to perform reduces or eliminates the age-difference in recall (Hess, Rosenberg, & Waters, 2001; Hess & Tate, 1991).

Research using daily diaries has shown that older adults use retrieval just as often as young adults in everyday tasks (i.e., preparing a meal, entering passwords to websites), which are often more familiar and regularly performed for older adults (Frank, Jordano, Brown, & Touron, 2016). This highlights the difference between memory usage by older adults in everyday life versus memory usage in laboratory tasks with novel information, and emphasizes the importance of familiarity and the bump in confidence that it brings. However, even in these familiar everyday tasks, older adults' willingness to use retrieval strategies was more strongly influenced by their memory confidence, whereas young adults used retrieval strategies based on their specific experiences with the task at hand. Confidence in one's memory abilities has been shown to be positively predictive of memory outcomes with older adults (Beaudoin & Desrichard, 2011; Berry, 1999).

Older adults' confidence in their memory, or perceived memory self-efficacy (MSE) (Bandura, 1989) has been a fertile ground of research in cognitive aging, as it is negatively correlated with increasing age and linked with memory performance (e.g., Beaudoin & Desrichard, 2011, 2017; Berry, 1999; Hertzog, McGuire, Horhota, & Jopp, 2010; Hertzog, McGuire, & Lineweaver, 1998; Lachman, Andreoletti, & Pearman, 2006; Light, 1991; Lineweaver & Hertzog, 1998), although the correlation between age and MSE is small (r = .15). Further, MSE has been shown to have a positive link between implementing effective encoding strategies (Hertzog et al., 1998; Lachman & Andreoletti, 2006; Lachman, Andreoletti, & Pearman, 2006; Riggs, Lachman, &

Wingfield, 1997; but see Hertzog et al., 2010), and so it is likely that MSE plays a role in strategic behavior in other tasks utilizing memory.

Cognitive Skill Acquisition

Cognitive skill acquisition tasks are useful for studying the interplay of metacognitive monitoring and control in older adults. A common feature of cognitive skill acquisition tasks is that they show a strategy shift over the course of the task from initially slow algorithmic processes (e.g., multistep rules that must be followed to reach the desired outcome) to later fast retrieval of the answer. Logan (1988) describes these two strategic processes as executed in parallel, where increasing the number of times or instances an item is encountered speeds up the retrieval process until it completes faster than the algorithm strategy. Rickard (1997; Bajic & Rickard, 2009, 2011) asserts that two strategies cannot be initiated simultaneously, so a strategy choice must initially be made. Given this strategy initiation bottleneck, some mechanism must exist that determines which strategy is selected (Bajic & Rickard, 2011).

One factor thought to influence strategy choice is a quick feeling-of-knowing (FOK) made when encountering an item (Lamson & Rogers, 2008; Reder & Ritter, 1992; Schunn, Reder, Nhouyvanisvong, Richards, & Stroffolino, 1997), defined here as an automatic judgment made upon seeing a cue of whether one knows the answer or not, based on increasing familiarity with the item over time. Reder and colleagues (Reder & Ritter, 1992; Schunn et al., 1997) used strategy choice as a proxy for FOK where participants had to choose on each trial of an arithmetic task either retrieve an answer to an arithmetic calculation or calculate the answer. This account argues that FOK is the governing process in strategy selection as opposed to a race model: that a strong or high FOK in the skill acquisition setting described above would lead one to retrieve whereas a weak or low FOK would lead one to use the algorithm.

Older adults shift to memory retrieval in these tasks later than young adults (Cerella, Onyper, & Hoyer, 2006; Rogers et al., 2000). One explanation is that older adults' associative deficit is solely to blame (Cerella et al., 2006), and thus the method of responding used (algorithm or memory) is driven by one's (slowly) accumulating knowledge of the task. Incorporating strategy choice into this account, older adults' associative deficit leads to slower acquisition of the items (and accompanying low FOKs) than young adults, resulting in longer periods of reporting the algorithm.

However, other work has suggested that the associative deficit does not account for all age differences in skill acquisition tasks, and that the later shifts to retrieval can also be explained as an avoidance of the retrieval strategy by older adults (Touron & Hertzog, 2004a, 2004b). In a task where a shift from algorithmic processing to retrieval-based processing was possible, Touron and Hertzog (2004a) periodically probed older adults' recognition memory throughout the task, finding that older adults' retrieval strategy choices lagged behind their measured knowledge of the task items, indicating older adults were opting to algorithmically verify the answer instead of just using their memory. Indeed, monetary incentives facilitate older adults' retrieval use without increasing their performance accuracy, response confidence, or cued recall accuracy, which were comparable to an older adult group offered no incentives (Touron, Swaim, & Hertzog, 2007). Further supporting the retrieval avoidance account over a pure associative deficit, Hertzog and Touron (2011) assessed FOK as a fast-deadline-constrained judgment prior to answering the question, and found that even with high FOKs, older adults were still reluctant to use retrieval. Even if older adults were confident that they knew the specific answer, they were still not confident in relying on retrieval outcomes.

To further investigate how the strategic set adopted by older adults differs from young adults as a function of knowledge of the task items, Hines, Hertzog, and Touron (2012) instituted a pre-learning condition into the task where participants learned either none, 50, or 100% of the task items. In the 100% pre-learning condition, young and old alike behaved as cognitive skill acquisition models would predict, quickly shifting to using retrieval (Cerella et al., 2006; Logan, 1988; Rickard, 1997). Older adults in the 50% condition opted instead to use the algorithm far more often for pre-learned items than the young adults in the same condition, where a race model would predict equivalent levels of retrieval usage between the two groups (i.e., retrieving for previously learned items, algorithmically processing new items). Hines et al. interpreted that the inclusion of the novel items in the 50% condition caused older adults to adopt a conservative strategy of responding using a consistent algorithm-based task set, even though they knew half of the items, again supporting a volitional account of strategy usage.

Moreover, theories of memory-based automaticity (Logan, 1988; Rickard, 1997) indicate that the shift to memory-based processing is only linked to the number of instances that the items have been experienced, with the speed of the memory-based process increasing per instantiation. Memory-based processing becomes the dominant processing mode when it becomes faster than algorithmic processing through practice, where exposure alone dictates the switch to memory (although Rickard's model does allow strategy choice, it is relatively agnostic as to the mechanism; Bajic & Rickard, 2011). As seen in the Touron et al. (2007) results, older adults in the incentives and non-incentives conditions experienced the same number of instances with task items, which contradicts mechanistic theories of instance-based responding that would predict responding by memory occurring at the same time for both groups. These patterns of results point to the retrieval strategy being under older adults' control, and not solely predicated on their slowly accumulating item knowledge.

Older adults do not always demonstrate retrieval reluctance in cognitive skill acquisition tasks, however. Rawson and Touron (2009, 2015) used a reading task where participants read novel noun-noun phrases with ambiguous meanings (i.e., skunk mud). After the phrase was used initially, it was later disambiguated using either the dominant meaning (i.e., most commonly given, e.g., mud as stinky as a skunk) or a subordinate meaning (not the most common, e.g., mud that had a dead skunk in it), as determined from pilot results where participants gave definitions of the noun-noun phrases. Shift to retrieval was determined by convergence of reading times between the two groups within age-group: reading time in the subordinate condition required reanalysis of the original meaning generated when first encountered, which was not necessary in the dominant condition. When the two groups' times converged, the subordinate group demonstrated no need to reanalyze the meaning of the term, simply retrieving the meaning from memory. Older adults fully converged upon the sixth repetition of the phrase (although later than the young), contrary to prior work showing only partial shifts to memory-based processing.

Rawson and Touron (2015) contrasted the shift to memory-based processing of older adults by having them also complete a standard skill acquisition task, alphabet arithmetic (verifying whether a presented letter was really a certain distance from another letter, e.g., F + 2 = H), where retrieval reluctance was found. The associative deficit does not easily account for these results, as unfamiliar noun pairings are usually not found to facilitate a shift to retrieval (Touron & Hertzog, 2004a, 2004b).

An important qualification of the Rawson and Touron (2009, 2015) findings is that reading does not offer a clear opportunity for strategy choice like other skill acquisition tasks, showing that older adults are sensitive to the task and task content in using their memory. Much work has been done in the domain of memory about the effects of task and task context, generally distinguishing between environmental support where the nature of the task provides a memory aid to older adults (Craik, 1983, 1986) and schematic support (Craik & Bosman, 1992), where features of the task items aid memory.

Environmental and Schematic Support

Craik (1983, 1986) viewed cognitive processing as including both processes that are driven by external stimulation, and those that are self-initiated. As the latter decline with increasing age, one becomes more dependent on external stimulation and environmental support to reduce deficits in performance. In memory tasks, recognition memory provides substantial environmental support as the target is re-presented.

However, in recall tasks few environmental cues may be present. Schematic support can be understood as use of prior knowledge to support richer encoding representations and to guide retrieval processes (Craik & Bosman, 1992). Craik and Jennings (1992) further predicted that varying levels of environmental and schematic support could have varied benefits for older adults' and young adults' associative memory.

Naveh-Benjamin, Craik, and Ben-Shaul (2002) investigated this prediction, considering the role of schematic and environmental support at encoding and retrieval. They showed participants pictures paired with sets of words that were related or unrelated to the picture, whereby the related words provided schematic support. In one experiment, two of the to-be recalled words had the first letter of the respective word given (environmental support), and the participant had to freely recall the last two words. For those words without environmental support, older adults' recall reflected standard age differences, with a benefit for related words. For those words with environmental support however, older adults' related-word recall was not statistically different from that of the young, reflecting a 180% jump in accuracy for older adults, and only a 13% increase for the young. Increases were more modest for unrelated words, 67% for older adults and 25% for young adults. Naveh-Benjamin et al. interpreted these results as showing that young adults can take advantage of beneficial encoding conditions (i.e., schematic support) above that of older adults, but with added support at retrieval (i.e., environmental support), older adults can greatly improve their performance, perhaps due to a reduced search set that the first letter of the to-be-recalled word affords.

Similar to these findings, Castel (2005) demonstrated that schematic support can eliminate age differences in associative memory. Castel paired pictures of grocery items with prices that varied by congruency: market-priced or unusual (under- and overpriced). Older adults recalled associated prices for market-priced items at a level equivalent to young adults, but showed lower recall for under- and overpriced items; this distinction did not affect recall in the young. Castel attributed the age-equivalence of the cued-recall of prices to schematic support, wherein prior knowledge facilitates processing of domain-relevant information (e.g., Ericsson & Kintsch, 1995; Hambrick & Engle, 2002; Van Overschelde, Rawson, Dunlosky, & Hunt, 2005; but see Umanath & Marsh, 2014 for situations where prior knowledge can be a detriment to accurate remembering).

The Present Study

The present study builds on the findings that schematic support can alleviate older adults' memory deficits, and incorporates it into an investigation of retrieval strategy use in a cognitive skill acquisition paradigm. The current work provided schematic support, with the expectation that this would increase older adults' willingness to use retrieval in a novel skill acquisition paradigm. Schematic support was expected to boost older adults' retrieval rates, as they should more efficiently learn schema-consistent information (Castel, 2005), and because easier-learned material boosts confidence that it will later be remembered above harder to learn information (Koriat, Ma'ayan, & Nussinson, 2006). Prior work has shown that easier items are learned faster than harder items for older adults, resulting in earlier shifts for those items, and later shifts to retrieval for harder items (Hoyer, Cerella, Onyper, 2003; Onyper, Hoyer, & Cerella, 2008). The differences in shifts to retrieval for the different item types has been explained as a mix of bottom-up and top-down processes – easy items don't require much effort to learn, however hard items can require conscious effort to learn, and older adults can have problems implementing top-down strategies in these cases (Hoyer et al., 2003; Onyper et al., 2008). The ease in learning that schematic support provides should influence older adults' strategy choice, such that the materials are capable of being retrieved sooner and the older adults feel more confident in using their memory.

The present study investigated the effects of schematic support on cognitive skill acquisition by using the Noun-Pair Lookup Task (NPLT), which is commonly used to investigate age-related differences in strategic memory use (Ackerman & Woltz, 1994; Rogers et al. 2000; Touron and Hertzog, 2004a, 2004b; Touron et al., 2007). The NPLT is a good tool to use in investigations of strategic memory use, due to ease with which the retrieval strategy can be isolated. Participants are tasked to verify on each trial whether a word pair is matched in a lookup table at the top of the screen, which is this task's algorithm. Half of the trials are matches, and half are mismatches. The word pairs are consistently mapped (e.g., Shiffrin & Schneider, 1977), in that the pairings never change and so can be learned over time. Additionally, the location of the word pairs in the lookup table changes every trial, so that a strategy dependent on their location is not developed. Removing the lookup table allows one to directly measure the participant's learning via recognition memory testing as the algorithm is then impossible to use to answer the question. If retrieval use in the task directly measures learning, percentage correct in recognition memory should roughly match the percentage of retrieval used in the block before. This is usually the case for young adults, however older adults present a different picture.

Work with the Noun-Pair Lookup Task has shown that there are many influences acting on older adults' motivation towards and confidence in memory strategy usage (Touron & Hertzog, 2004a, 2004b; Touron et al., 2007). These include the perceived affordability of strategy shift, demonstrated by an interaction between number of items to learn and difficulty in scanning the table, where a small number of items coupled with greater scanning difficulty will cause older adults to shift to retrieval at a faster rate (Touron & Hertzog, 2004a). Periodically testing memory of the item pairs also appears to improve memory confidence, showing higher levels memory use by older adults who (successfully) complete memory tests relative to those who are not tested on their memory (Touron & Hertzog, 2004a).

The present study also used a novel variation of the Noun-Pair Lookup Task that includes materials that are closely related to what Castel (2005) used: grocery items paired with prices, either market-priced or overpriced. The task (an isomorph of the NPLT, hereby referred to as the Grocery-Pair Lookup Task) pairs 12 grocery items with prices (e.g., Eggs – $2.79, Milk – $1.89). For each trial, a pair will appear that either is a direct match with the 12 pairs presented in a table directly above, or a mismatch (e.g., Eggs – $1.89). With enough exposure participants will learn the pairs and begin to answer with just their memory, without having to scan the table for the answer. Strategy reports were given after each trial, indicating use of the algorithm, memory, both, or some other strategy. By providing schematic support in the form of market-based prices, older adults are expected to use memory-based processing sooner and more often than those in the overpriced condition, due to the increased confidence in memory use that the market-priced condition affords.

Goals and Hypotheses

Older and young adults completed the Noun-Pair Lookup Task and then either a market-priced or overpriced version of the Grocery-Pair Lookup Task. I predicted that older adults in the market-priced condition will shift to using their memory sooner than older adults in the overpriced condition, as indicated by their strategy reports given after each trial. This finding will provide an important extension of the finding of the selective preservation of memory-based automaticity in older adults (Rawson & Touron, 2009; 2015), and inform theories of skill acquisition by clarifying the factors that affect strategy choice and situations in which older adults feel confident in memory use. Critically, I predicted that older adults' retrieval use would not differ by condition in the Noun-Pair Lookup Task, emphasizing the importance of schematic support on memory use / avoidance. Because of the extensive practice with the material, older adults' memory use in the Noun-Pair Lookup Task and in the overpriced condition of the Grocery-Pair Lookup Task was expected to largely differ from their memory ability, i.e., what they know as assessed by the recognition memory test, implicating an avoidance of memory rather than a lack of memory. The young adults were expected to perform similarly in both tasks, as memory in young as not been shown to differ with the market/overpriced distinction (Castel, 2005), and manipulations shown to alter older adults' retrieval strategy choice have not been shown to impact young adults (e.g., Touron, 2015).

Post-task measures of task confidence, including global confidence in the task (i.e., confidence in using the memory strategy) and estimated number of memorized items, were expected to correlate with end-of-task retrieval usage for older adults, replicating previous findings (Touron & Hertzog, 2004b; Touron et al., 2007). This would support the contention that confidence is an important factor in strategy choice in skill acquisition. Finally, I correlated older adults' task confidence ratings and end-of-task retrieval usage with their beliefs about memory as assessed through the Personal Beliefs About Memory Instrument (PBMI), which measures perceived changes in memory and memory use throughout the lifespan and self-referent memory beliefs, including memory self-efficacy (Lineweaver & Hertzog, 1998; Hertzog, Lineweaver & Hines, 2014). There is evidence to expect a positive correlation, as found in an individual-differences study of strategy-use in the NPLT (Touron, 2015), but null findings have also been obtained with smaller sample sizes (Touron & Hertzog, 2004a) and correlations of performance and memory self-efficacy are often small (Beaudoin & Desrichard, 2011).

CHAPTER II

METHOD

Participants

One-hundred and forty subjects were collected in total, but only 120 were kept for analyses as it was the planned *N* (12 young adult subjects collected after planned *N* of 120 was met[1]; four subjects excluded for program failure – three young adults: two in the market-priced condition, one in the overpriced condition, and one older adult in the overpriced condition; two subjects for being outside normal age range of 18 to 35 for young adults (both over 45 years old, one each in the market- and overpriced conditions); one young adult in the overpriced condition excluded for noncompliance with instructions; one older adult in the overpriced condition voluntarily withdrew from study). All participants that were excluded before the planned *N* was met were later replaced. Young adults were run in the spring semester of 2017, and were compensated with course credit. Older adults were run from May to October 2017, and were compensated with $30.

For the participant characteristics, typical age differences were obtained, which are shown in Table 1. Older adults in this sample were more educated than the young

[1] There were no significant differences in participant characteristics between those collected after the planned *N* and those collected before the planned *N* was met, *p*s > .13.

adults[2], $F(1, 115) = 102.11$, $MS_{Error} = 3.34$, $p < .001$, $\eta^2_G = .47$. Older adults also took more medications than young adults, $F(1, 115) = 28.04$, $MS_{Error} = 1.88$, $p < .0001$, $\eta^2_G = 0.19$. Older adults scored higher on the vocabulary measure (Shipley Vocab), a measure of crystallized intelligence, $F(1, 116) = 84.69$, $MS_{Error} = 16.62$, $p < .0001$, $\eta^2_G = .42$. Older adults performed worse at the digit-symbol substitution task, (a measure of processing speed), $F(1, 116) = 28.09$, $MS_{Error} = 100.8$, $p < .0001$, $\eta^2_G = .19$. For digit symbol recall, older adults recalled fewer symbols, $F(1, 116) = 51.35$, $MS_{Error} = 3.65$, $p < .0001$, $\eta^2_G = .31$, ($M = 7.35$, $SE = 0.25$; $M = 4.85$, $SE = 0.25$). Unexpectedly, a condition difference was found for digit symbol recall, $F(1, 116) = 4.42$, $MS_{Error} = 3.65$, $p = 0.037$, $\eta^2_G = .036$, which appeared to be driven by young adults in the overpriced group recalling fewer symbols than young adults in the market-priced group ($M = 6.77$, $SE = 0.35$; $M = 7.93$, $SE = 0.35$, respectively). However, post-hoc comparisons (Tukey's HSD) revealed that young adults' symbol recall in the two conditions were not significantly different from each other, $p = .089$, $d = 0.61$. There was also no condition difference between older adult groups, $p = .93$, $d = 0.16$.

Design

Two between-subjects conditions were used – age (young, old) and Grocery-Pair Lookup Task condition (market-priced, overpriced). Equal numbers of participants were randomly assigned to each condition. Practice block (1 – 20 for training, 1 – 2 for

[2] Data for one older adult were missing for education and medication measures due to an experimenter-based error.

recognition memory) was the within-subjects manipulation. The dependent variables include task accuracy, retrieval use, reaction time, and recognition memory performance.

Materials

Visual Basic 6.0 and E-Prime 2.0 were used to present the tasks and questionnaires to participants. Stimuli were presented in Arial font on a 15-in (38.1 cm) LCD monitor with a resolution of 1024 x 768. The stimulus set for the Noun-Pair Lookup Task (NPLT) contained 12 unrelated concrete noun pairs (i.e., ivy-bird, potato-frog). All 12 noun pairs were presented in the lookup table for standard trials; the location for each pair changes within the table each trial, but the pairings were consistent. A central pair was matched to one of the pairs within the table for half of the presentations, and for the other half a mismatched pair was presented (i.e., ivy-frog). On each trial participants verified whether the centrally-presented pair is matched in the table (by pressing keys labeled “Y” for a match, “N” for a mismatch”), and then a separate screen asks them to report the strategy they used to respond. Due to these features, the task initially requires visual search to be successful, but with increasing trials the pairings may be learned and participants can rely on their memory. Strategy-report options include keys labeled “S” for scanning, “M” for retrieving the answer from memory, “B” for a combination of both scanning and memory, or “O” for other.

The Grocery-Pair Lookup Task (GPLT) is an isomorph of the NPLT: 12 grocery items (accompanied by corresponding pictures) and prices are presented in the lookup table during a standard trial, as shown in Figure 1. The locations of the 12 item-price pairs change every trial, but each pairing remains consistent, allowing the pairings to be memorized with repetitions. Half of the trials were an exact match of a pairing from the table; the other half were mismatched, with a random price paired with a random item (the pictures of each item will always be consistent even if the pairings aren't). Market-priced item prices are between \$1 and \$2.99, with all prices ending in 9s. Overpriced item-prices are between \$12 and \$13.99, again with all prices ending in 9. Prices are constrained so that when re-paired during mismatch trials, the paired price will make plausible sense (or approximately equal implausible sense for overpriced items), which is particularly important for the market-priced condition to maintain consistent schematic support (e.g., Soda – \$4.99 could quickly be distinguished as unusual, whereas Soda – \$2.99 is plausible). I constrained the prices within a condition so they would not be sufficiently distinctive that familiarity-based memory rather than more detailed recollection could be used to respond.

Procedure

Participants were run in small groups (from one to four) on computers separated by dividers, with the experimenter present. Participants in each session were assigned to the same condition, where condition was assigned so that approximately equal numbers of participants were in each condition as the study was active. Participants first gave informed consent, and then completed a questionnaire requiring basic demographic information about their age, self-reported health, years of education, and daily medication use. Next, they completed the cognitive pre-tests described in Table 1. Participants then completed the NPLT, starting with practice trials and then 20 blocks of 24 trials each. Following training, participants received two blocks of recognition memory probes, where the matched and mismatched stimuli were presented without the lookup table; participants were not initially informed of these memory tests. Each block in the task was followed with a mandatory ten-second break and the option to take a longer break if necessary, to prevent fatigue.

Following the NPLT, participants completed a survey regarding their previous task performance, including overall task confidence, estimated number pairs learned, overall memory-use confidence, item-level cued judgments of learning (i.e., showing the left-hand pair and asking for confidence, 0-100, that they will remember the right-hand pair), and cued-recall for the right-hand pair.

Participants next completed the GPLT, assigned to either the market-priced or overpriced condition. The task is functionally similar to the NPLT, with 20 blocks of standard trials with strategy probes followed by two blocks of recognition memory probes. The condition difference comes down to the prices shown, either market-priced or overpriced. Participants again completed several practice trials before beginning the task, with option to repeat the instructions and practice, and again were not informed of the later memory tests.

The task was also followed by a survey regarding the previous task performance, including the same kind of confidence questions asked previously for the NPLT: estimated number of pairs learned, cued JOLs, and cued-recall, in addition to questions about overall self-evaluated study performance and confidence, and ratings of task difficulty, comparing the NPLT with the GPLT.

Participants finally completed a questionnaire about memory beliefs, the General and Personal Beliefs about Memory Instrument (GBMI and PBMI) (Hertzog & Lineweaver, 1998; Hertzog et al., 2014). The GBMI measures general beliefs about how memory changes over the lifespan. Data from the GBMI are not reported here because they were not theoretically relevant. The PBMI measures memory self-concept, perceived and expected memory change, and perceived control over memory. It took young adults about two hours to complete the experiment, whereas older adults took up to three hours.

CHAPTER III

RESULTS

For all analyses, α was set at .05, unless otherwise specified. All hypotheses are two-tailed against a null effect unless otherwise specified. Two-sided Welch's *t*-tests were used for *t*-statistics, except in post-hoc tests where Tukey HSD tests were run. Welch's *t*-test was preferred as it provides better control over Type I error rates than the Student's *t*-test while losing at most ~6% in power under most circumstances (see Delacre, Lakens, & Leys, 2017). Generalized eta-squared will be reported for ANOVA effect sizes (Olejnik & Algina, 2003). Generalized eta-squared is similar to partial eta-squared, but differs in that the variance from observed between-subjects differences (e.g., age, gender) is considered in the denominator for each predictor, which improves the comparison of effect sizes between studies that don't share identical designs, and is particularly recommended for studies with repeated measures (Bakeman, 2005). ANOVAs and post-hoc tests were run using the *afex* package (Singmann, Bolker, Westfall, & Aust, 2017) in the *R* statistical programming language (R Core Team, 2017).

Condition (i.e., assignment to the market-priced or overpriced version of the GPLT) was included in all analyses as a between-subjects condition (and its interaction with age), although not always theoretically motivated (i.e., prior to the experimental manipulation), intended as a check on my manipulation to rule out individual differences. Repeated measures analyses were adjusted for Greenhouse-Geisser corrections. For sake of brevity in reporting specific contrasts, young adults in the market-priced condition will be referred to as YA-MP, young adults in the overpriced condition will be referred to as YA-OP, older adults in the market-priced condition as OA-MP, and older adults in the overpriced condition as OA-OP.

Analyses will be presented alongside my initial predictions. Demographic and pre-test measure statistics were presented in the above Participants section. Below, analyses are first presented for the GPLT, examining differences in task accuracy, retrieval strategy use, reaction times, and recognition memory in turn. Equivalent analyses for the NPLT are then presented. Correlations of retrieval use and confidence in retrieval use between tasks are reported next. Then, post-task questionnaire data are presented, including correlations of post-task questionnaire items with retrieval use at the end of practice with the GPLT and NPLT. Finally, results from the PBMI are briefly noted. Tables and figures are presented in Appendix A. The lists of questions used in the post-task questionnaires are presented in Appendix B. The full set of analyses with the PBMI are presented in Appendix C, as this measure was supplementary to the goals of this project.

Grocery-Pair Lookup Task

Task accuracy. Two participants were removed from all GPLT analyses and subsequent cross-task analyses (i.e., with the NPLT) for having accuracies not distinguishable from chance by binomial test (both p-values were less than an alpha set to .0004, or .05/120; one young adult each from the market-priced and overpriced groups, Ms = .53, and .57, respectively). Although the assumption of independence for the binomial test is certainly violated in this sample, as performance on each trial is not independent from other trials which could increase the probability of a false positive, the strict alpha adopted here and inspection of the participants' response patterns (i.e., long strings of responding only 'Yes' or 'No') increase my confidence in rejecting these participants.

The accuracy data for the GPLT revealed a complex and not easily interpretable pattern. On the whole, older adults were more accurate than young adults, (M = .91, SE = .01; M = .96, SE = .001, for young and older adults, respectively), $F(1, 114) = 19.31$, $MS_{Error} = 0.075$, $p < 0.001$, $\eta^2_G = 0.085$. Older adults have been found to be more accurate in the NPLT in previous studies (Touron & Hertzog, 2004a; 2004b; although not always, cf. Touron et al., 2007), so this result is not surprising. This was qualified by an age by condition interaction, $F(1, 114) = 5.31$, $MS_{Error} = 0.075$, $p = 0.023$, $\eta^2_G = .023$. Young adults in the overpriced, but not market-priced condition were less accurate overall than older adults, (M = .93, SE = .01; M = .89, SE = .01; M = .96, SE = .01, M = .97, SE = .01, for YA-MP, YA-OP, OA-MP, and OA-OP, respectively), $t_{crit} = 2.537$: YA-OP vs YA-MP: $t(114) = 2.50$, $p = .065$, $d = 0.15$; YA-OP vs OA-MP: $t(114) = 4.002$, $p < .001$, $d = 0.23$; YA-OP vs OA-OP: $t(114) = 4.74$, $p < .0001$, $d = 0.28$; YA-MP vs OA-MP: $t(114) = 1.48$, $p = .45$, $d = 0.09$; YA-MP vs OA-OP: $t(114) = 2.21$, $p = .13$, $d = 0.13$; OA-MP vs OA-OP: $t(114) = .74$, $p = .88$, $d = 0.04$. There was a main effect of block, $F(10.75, 1225.14) = 2.58$, $MS_{Error} = .005$, $p = 0.003$, $\eta^2_G = 0.008$, qualified by a three-way interaction of age by condition by block interaction, $F(10.75, 1225.14) = 2.38$, MS_{Error} = .075, $p = 0.009$, $\eta^2_G = 0.008$.

It appears that the younger adult overpriced group became less accurate over time, from a first block accuracy of $M = .94$, $SE = .01$, to their final block accuracy of $M = 0.86$, $SE = 0.03$, a difference of .08, where all other groups experienced changes $\leq .03$ (YA-MP: first block: $M = .94$, $SE = .02$; final block: $M = 0.95$, $SE = 0.01$; OA-MP: first block: $M = .98$, $SE = .005$; final block: $M = .96$, $SE = .01$; OA-OP: first block: $M = .97$, $SE = .01$; final block: $M = .97$, $SE = .01$). A trend analysis for accuracy was tested, to attempt to reveal any systematic effects of block; presence of significant higher order trends above the cubic would be interpreted as noise. For the three way interaction of age by condition by block, higher order trends were significant at the cubic, $F(1, 114) = 8.41$, $MS_{Error} = 0.003$, $p = .004$; 11th order, $F(1, 114) = 6.12$, $MS_{Error} = 0.003$, $p = .015$; 15th order, $F(1, 114) = 4.011$, $MS_{Error} = 0.002$, $p = .048$; and 18th order, $F(1, 114) = 5.66$, $MS_{Error} = .003$, $p = .019$. These fluctuations are not of theoretical interest and will not be discussed further. Overall, accuracy was high across conditions and blocks.

Retrieval use. Retrieval use increased with practice, shown in Figure 2. Young adults retrieved more than older adults, the market-priced group retrieved more than overpriced group, and the market-priced group increased their rates of retrieval faster than the overpriced group. These results were confirmed, with main effects of age, $F(1, 114) = 18.69$, $MS_{Error} = 1.00$, $p < 0.0001$, $\eta^2_G = .094$, condition, $F(1, 114) = 46.22$, $MS_{Error} = 1.00$, $p < 0.0001$, $\eta^2_G = .189$, and block, $F(4.87, 554.85) = 102.97$, $MS_{Error} = 0.12$, $p < 0.0001$, $\eta^2_G = .226$. A condition by block interaction was significant, $F(4.87, 554.85) = 6.04$, $MS_{Error} = 0.12$, $p < 0.0001$, $\eta^2_G = .017$.

By the final block of practice, older adults in the market-priced condition retrieved as much as young adults in the same condition, but were also indistinguishable from young adults in the overpriced condition (YA-MP: $M = .85$, $SE = .06$; OA-MP: $M = .75$, $SE = .06$; YA-OP: $M = .62$, $SE = .06$; OA-OP: $M = .39$, $SE = .06$). Older adults in the overpriced condition retrieved less than each other group. This was confirmed by a two-way ANOVA of condition by age in the last block of practice, with main effects of age, $F(1, 114) = 8.54$, $MS_{Error} = .093$, $p = .004$, $\eta^2_G = .069$, and condition, $F(1, 114) = 26.26$, $MS_{Error} = .093$, $p < 0.0001$, $\eta^2_G = .175$. The condition by group interaction was not significant, $p = .25$. Tukey's HSD tests showed that young and older adults in the market-priced condition were not significantly different, $p = .59$, $d = 0.32$, YA-MPs retrieved more than OA-OPs, $p < .001$, $d = 1.48$, OA-MPs were not significantly different from YA-OPs, $p = .4$, $d = 0.41$, YA-MPs retrieved more than YA-OPs, $p = .03$, $d = .73$, YA-OPs retrieved more than OA-OPs, $p = 0.024$, $d = .75$, and OA-MPs retrieved more than OA-OPs, $p = .0001$, $d = 1.16$.

The pattern of retrieval use that was observed was not exactly as predicted. My prediction that older adults in the market-priced group would retrieve more than older adults in the overpriced group was confirmed. Unexpectedly, young adults also showed differences in retrieval use depending on condition, where those in the market-priced condition retrieved more than those in the overpriced condition. Condition differences between young adults are not usually found, as previous work with the NPLT has shown (e.g., Touron & Hertzog, 2004b; Touron et al., 2007; cf. Touron & Hertzog, 2004b).

These results suggest that schematic support can alter retrieval use in both young and older adults.

Reaction times. Reaction times sped up with increasing practice. Main effects of age, $F(1, 114) = 145.55$, $MS_{Error} = 4{,}312{,}483$, $p < .0001$, $\eta^2_G = .443$, condition, $F(1, 114) = 28.57$, $MS_{Error} = 4{,}312{,}483$, $p < .0001$, $\eta^2_G = .079$, and block, $F(7.48, 852.57) = 116.75$, $MS_{Error} = 328365$, $p < .0001$, $\eta^2_G = .168$, were found. These effects are qualified by age by block, $F(7.48, 852.57) = 3.80$, $MS_{Error} = 328{,}365$, $p = .0003$, $\eta^2_G = .007$, and condition by block interactions, $F(7.48, 852.57) = 3.04$, $MS_{Error} = 328{,}365$, $p = .003$, $\eta^2_G = .005$. See Figure 3 for reaction time data plotted out by strategy. Unsurprisingly, younger adults were faster than older adults, and increased their reaction times at a faster pace. The market-priced group was faster than the overpriced group, and increased their reaction times at a faster rate. This pattern is likely primarily due to differences in strategy usage – retrieving is faster than scanning.

Strategy report validation. Strategy reports in shift-to-retrieval tasks such as those used in this study have been validated before, and among different kinds of tasks (Rickard, 1997; Touron et al., 2004, Touron & Hertzog, 2004b). However, because this was a new task, I felt that validating the strategy reports was warranted. To validate the use of strategy reports, reaction time distributions were split between retrieval and scanning across the whole task, collapsing across age-groups and condition, where retrieval reaction times are expected to be faster than scanning reaction times. Reaction times were rank-ordered within participants into deciles (i.e., 10 bins), from which the reaction times were averaged across individuals and then decile, see Figure 4. Cumulative distribution functions (CDFs) for both strategies were created from these data (Miller, 1982; Nino & Rickard, 2003; Townsend, 1990; Van Zandt, 2000). No cross-over between the distributions when examining their CDFs indicates that the two distributions are distinct from one another (Townsend, 1990). A cross-over in distribution functions would call in to question the validity of the strategy reports, particularly at the fastest ends of the distributions.

Paired t-tests were run to compare mean reaction time at each decile (the data are normally distributed when separated like this, Van Zandt, 2002). I set alpha to .005 to correct for multiple comparisons. Participants with decile cell sizes fewer than 10 (i.e., at least 10 trials) were removed from the analyses (removing 52 participants in the scanning distribution, 15 in the retrieving distribution), and participants not present in both the scanning and retrieving data sets were removed (final N = 44, $df = 43$). All points of comparison between the distributions were significantly different, all *p*s < .005, Cohen's d_z from 1.23 to 1.91 (where $d_z = t / \sqrt{n}$, e.g., Lakens, 2013), mean differences in RTs ranging from 535 ms to 1624 ms. A two-sample Kolmogorov-Smirnov test was run on the distribution functions themselves, a nonparametric test to tell if the two functions are different, which was the case, $D = 0.7$, $p = .01$. These analyses confirm the validity of the strategy reports.

Recognition memory. Following the training trials, participants completed recognition memory tests that presented the same central target pairs but without the lookup table. The purpose of these recognition memory tests is to distinguish memory strategy use in training from memory strategy ability. I did not have any predictions regarding performance given the novelty of the task. I calculated *d'* (z(hits) – z(false alarms)) to compare recognition memory between groups. Hit and false alarm rates were corrected when at ceiling or floor, such that either 1 - 1/(2N) or 1/(2N) were used, for ceiling and floor rates, where N = number of trials, here equal to 12. A 2 (age) by 2 (condition) by 2 (block) repeated measures ANOVA was conducted, revealing a significant effect of condition, $F(1, 114) = 51.65$, $MS_{Error} = 1.26$, $p < .0001$, $\eta^2_G = .27$, where the market-priced group scored significantly higher than the overpriced group ($M = 2.95$, $SE = 0.1$; $M = 1.90$, $SE = 0.1$, for market-priced and overpriced groups, respectively). All other effects were not significant, *p*s > .15.

Estimating retrieval reluctance. To the extent that retrieval use and recognition memory are both indexing learning, I ran a repeated measures ANOVA, predicting retrieval use in the final block of practice and accuracy over the two recognition memory tests (both on the same scale, as proportions) with age, condition, and block as predictors. Significant differences between the last block of practice and the recognition memory test can be interpreted as an index of retrieval reluctance. I found evidence of retrieval reluctance in this sample. There were main effects of age, $F(1, 114) = 6.63$, $MS_{Error} = .068$, $p = .01$, $\eta^2_G = .03$; condition, $F(1, 114) = 46.78$, $MS_{Error} = .068$, $p < .0001$, $\eta^2_G = .18$; and block, $F(1.17, 133.29) = 68.31$, $MS_{Error} = .046$, $p < .0001$, $\eta^2_G = .19$. There were also significant interactions of age and block, $F(1.17, 133.29) = 7.10$, $MS_{Error} = .046$, $p = .006$, $\eta^2_G = .03$, and condition and block, $F(1.17, 133.29) = 8.05$, $MS_{Error} = .046$, $p = .003$, $\eta^2_G = .03$. Post-hoc focused comparisons help to reveal the underlying pattern.

All groups except for the YA-MP group had significant differences between the last block of practice and the first block of recognition memory: YA-MP: $t(228) = 2.44$, $p = .38$, $d = .49$; OA-MP: $t(228) = 4.53$, $p = .0006$, $d = .89$; YA-OP: $t(228) = 4.51$, $p = .006$, $d = .91$; OA-OP: $t(228) = 8.79$, $p < .0001$, $d = 1.74$. No differences were found between the first and second blocks of recognition memory performance, as previously reported. This pattern of results suggests that all groups except for YA-MP were underutilizing the retrieval strategy in the final block of practice. To the extent that retrieval use and recognition memory are not totally commensurate indices, I next report a regression model using recognition memory as a covariate in predicting retrieval use.

Retrieval use as a function of recognition memory. If it were the case that retrieval use in the final block of practice merely reflected pair learning, covarying recognition memory performance (i.e., *d'*) should remove any effects of group and age (using proportion accuracy produces the same pattern of results). I ran a regression predicting retrieval use in the final block of practice with age, condition, and *d'* as predictors. This regression is plotted in Figure 5. After adding *d'* into the model, a main effect of age was observed, $F(1, 110) = 5.08$, $MS_{Error} = .072$, $p = .026$, $\eta^2_G = .04$, where young adults retrieved more than older adults, along with an effect of *d'*, $F(1, 110) = 19.99$, $MS_{Error} = .072$, $p < .001$, $\eta^2_G = .14$, where a higher score implied greater retrieval. Condition was not significant, $F(1, 110) = 0.46$, $MS_{Error} = .072$, $p = .49$, $\eta^2_G = .004$, nor were any two-way interactions, *p*s > .06. The three-way interaction between age, condition, and *d'* was significant, however, $F(1, 110) = 4.53$, $p = .036$, $MS_{Error} = .072$, $\eta^2_G = .036$. The three-way interaction indicates differing slopes between each age and condition combination. I can conclude that *d'* predicted different rates of retrieval between each age and condition combination, i.e., that retrieval ability did not totally predict retrieval use.

Noun-Pair Lookup Task

Repeated measures ANOVAs will be reported below, using a 2 (age) by 2 (condition) by 20 (block) design unless otherwise noted, testing for all higher order interactions. As a reminder, participants first completed the NPLT, and Noun-Pair post-task questionnaire before completing the GPLT.

Task accuracy. Two participants were removed from analyses due to accuracy statistically indistinguishable from chance by binomial test, alpha set to .0004, $M = .51$, $M = .5$, respectively (one young adult from the overpriced condition, and one older adult from the overpriced condition). These participants were also removed from any cross-task analyses. As mentioned previously, the probability of type I error is increased due to dependence of responding between trials, the strict alpha threshold and long strings of repeated Yes or No decisions by these participants increase my confidence in this decision (the older adult rejected here responded only “No” to all trials).

Older adults were more accurate than young adults ($M = .93$, $SE = .01$; $M = .97$, $SE = .01$, for young and older adults, respectively), $F(1, 113) = 25.43$, $MS_{Error} = .033$, $p < .0001$, $\eta^2_G = .08$. No other effects were significant. This pattern of results is typical of performance in the NPLT (e.g., Touron & Hertzog, 2004a, 2004b; Touron et al., 2007).

Retrieval use. Young adults used retrieval more often than older adults, shown in Figure 6, $F(1, 112) = 95.84$, $MS_{Error} = 1.04$, $p < .0001$, $\eta^2_G = .35$. Retrieval use increased over blocks, $F(4.59, 514.69) = 107.34$, $MS_{Error} = 0.118$, $p < .0001$, $\eta^2_G = .17$. Young adults also increased their retrieval use at a faster rate than older adults, indicating an age by block interaction, $F(4.59, 514.69) = 16.62$, $MS_{Error} = 0.118$, $p < .0001$, $\eta^2_G = .03$. No other effects were significant.

In the final block practice, young adults retrieved more than older adults ($M = .90$, $SE = .05$; $M = .82$, $SE = .05$; for young adults in the market-priced and overpriced conditions, respectively; $M = .49$, $SE = .05$; $M = .32$, $SE = .05$, for older adults in the market-priced and overpriced groups, respectively), $F(1, 112) = 73.88$, $MS_{Error} = .083$, $p < .0001$, $\eta^2_G = .39$, and unexpectedly, an effect of condition, $F(1, 112) = 5.64$, $MS_{Error} = .083$, $p = .019$, $\eta^2_G = .03$. This was unexpected as there was no manipulation of condition for the NPLT, which was completed before the GPLT. Inspection of Figure 6 suggests that participants in the overpriced group were retrieving less than those in the market-priced group, perhaps responsible for the significant finding above. Post-hoc analyses showed no significant differences between the two older adult groups, $p = .09$, $d = .62$, or young adult groups, $p = .74$, $d = .27$. These results may indicate that participants in the overpriced condition, by chance, had a lesser tendency to transition to retrieval strategies.

Reaction times. On the whole, young adults were faster than older adults throughout the task, $F(1, 113) = 172.98$, $MS_{Error} = 11{,}454{,}105$, $p < .0001$, $\eta^2_G = .52$. Reaction times sped up with practice, revealed by a main effect of block, $F(4.95, 558.82) = 164.59$, $MS_{Error} = 902{,}930$, $p < .0001$, $\eta^2_G = .16$. An age by block interaction was significant, $F(4.95, 558.82) = 2.25$, $MS_{Error} = 902{,}930$, $p = .048$, $\eta^2_G = .003$, where young adults improved at a faster rate than older adults. There were no other significant effects or interactions, *p*s > .08.

Strategy report validation. Although strategy reports have been validated in the Noun-Pair Lookup Task before (Touron & Hertzog, 2004a, 2004b; Touron et al., 2007), nonetheless I compared the separate cumulative distribution functions for scanning and retrieval as with the Grocery-Pair data. Reaction times were rank ordered into deciles within each participant, averaged by decile and then averaged across all participants. I set alpha to .005 to correct for multiple comparisons, as 10 paired *t*-tests were run. Participants with decile cell sizes fewer than 10 (i.e., at least 10 trials) were removed from the analyses, and participants not present in both the scanning and retrieving data sets were removed (final N = 29). All points of comparison between the distributions were significantly different, all *p*s < .005, d_z from 1.73 to 2.97, mean differences in RTs ranging from 690 ms to 2767 ms. A two-sample Kolmogorov-Smirnov test was run on the distribution functions, $D = 0.8$, $p = .002$, indicating separation of the two functions. These results once more confirm the validity of the strategy reports.

Recognition memory. I calculated *d'* for recognition memory, comparing scores across the two blocks of recognition memory tests. Hit and false alarm rates were corrected for performance at ceiling and floor, such that either 1 - 1/(2N) or 1/(2N) were used, for ceiling and floor rates, respectively. Two participants were removed from the analyses for scoring at chance due to only responding Yes or No (two older adults from the market-priced, and overpriced conditions, respectively). Young adults were more accurate than older adults ($M = 2.97$, $SE = .1$; $M = 2.00$, $SE = .1$, for young and older adults, respectively), yielding an effect of age, $F(1, 111) = 38.92$, $MS_{Error} = 1.36$, $p < .0001$, $\eta^2_G = .24$. Block was not significant, $F(1, 111) = 0.11$, $MS_{Error} = 0.17$, nor were any other effects, *p*s > .09. Age differences in recognition memory have been found routinely (Touron & Hertzog, 2004b, 2009; Touron et al., 2007). Improvements in memory performance across blocks has also been observed (Touron & Hertzog, 2009; Touron et al., 2007), and it is unclear why it did not occur in the present study.

Estimating retrieval reluctance. I ran the equivalent analyses reported above for the NPLT, predicting retrieval use in the final block of practice and recognition memory accuracy (both on the same scale, as proportions) with age, condition, and block as predictors. Significant differences between the last block of practice and recognition memory can be taken as an index of retrieval reluctance. An unexpected pattern of results emerged. There were main effects of age, $F(1, 113) = 66.25$, $p < .0001$, $MS_{Error} = .083$, $\eta^2_G = .26$; condition, $F(1, 113) = 4.10$, $MS_{Error} = .083$, $p = .045$, $\eta^2_G = .02$; and block, $F(1.19, 134.38) = 68.31$, $MS_{Error} = .029$, $p < .0001$, $\eta^2_G = .17$. There were significant interactions of age and block, $F(1.19, 134.38) = 59.86$, $MS_{Error} = .029$, $p < .0001$, $\eta^2_G = .09$, and condition and block, $F(1.19, 134.38) = 3.99$, $MS_{Error} = .029$, $p = .04$, $\eta^2_G = .006$. Focused comparisons confirmed my expected results, despite unexpected condition differences.

Only older adults had significant differences between the last block of practice and the first block of recognition memory: YA-MP: $t(226) = 1.54$, $p = .92$, $d = .31$; OA-MP: $t(226) = 9.63$, $p < .0001$, $d = 1.95$; YA-OP: $t(226) = 2.65$, $p = .26$, $d = .54$; OA-OP: $t(226) = 13.18$, $p < .0001$, $d = 2.67$. This pattern of results conforms to the expectation that older adults were underutilizing retrieval. The significant effect of condition, and its interaction with age seem show that the overpriced condition tended to have larger disparities between retrieval use and recognition memory accuracy than the market-priced condition, although this pattern didn't hold for adjustment to multiple comparisons. The unexpected condition difference was likely due to older adults in the overpriced condition retrieving less than those in the market-priced condition, although young adults in the overpriced condition also retrieved numerically, but not statistically less than those in the market-priced condition.

Retrieval use as a function of recognition memory. Recognition memory performance (*d'*) was entered in as a continuous covariate in the regression of retrieval use in the final block of practice on age and condition, and an interaction with age and condition. The regression is plotted in Figure 7. There was no significant interaction of *d'* with age or group, *p*s > .30. Dropping those interactions with the model caused no significant decrease in model fit, $F(1, 106) = 0.48$, $p = .69$. Dropping the interaction of age and condition also caused no significant decrease in model fit, $F(1, 109) = 1.15$, $p = .28$. Dropping condition also caused no significant decrease in model fit, $F(1, 110) = 1.99$, $p = .16$, however dropping age from the model with condition and *d'* did cause a decrease in model fit, $F(1, 110) = 28.39$, $p < .0001$, so age was kept in the model and condition was dropped. Dropping either age, $F(1, 111) = 27.32$, $p < .0001$, or *d'*, $F(1, 111) = 111.96$, p < .0001, from the model provided significantly worse fit than the ANCOVA model of age and *d'*. In this model, effects of *d'* and age were significant, $F(1, 111) = 111.96$, $MS_{Error} = .044$, $p < .0001$, $\eta^2_G = .45$, and $F(1, 111) = 27.32$, $MS_{Error} = .044$,

$p < .0001$, $\eta^2_G = .19$, respectively. This replicates prior work (Touron & Hertzog, 2004b) showing that learning reflected in recognition memory tests is not totally responsible for retrieval use, and that effects of age explain lower rates of retrieval use as well.

Retrieval use between tasks. I correlated retrieval use in the final block of the Noun-Pair and Grocery-Pair Lookup Tasks, to detect if patterns of retrieval use were stable across the two tasks, collapsed across age-group and condition to increase power. This relationship is plotted in Figure 8. Pearson and Spearman correlations were calculated, as these data were highly skewed, in order to determine the linear relationship of retrieval between the two tasks, and to determine if the ranked order of proportion retrieval was preserved across tasks, respectively. Additionally, participants that never endorsed retrieving were excluded from these analyses.[3] The two correlation coefficients returned similar results.[4] The suite of correlations are reported in Table 2. The correlations were significant when including the whole sample, $r = .29$, $p = .002$, 95% CI [.11, .46]; $r_s = .25$, $p = .01$, 95% CI [.06, .42]. These results indicate that there is a small-to-modest linear and increasing monotonic relationship in retrieval use between the two tasks.

Retrieval use was calculated between both age groups and both conditions, and then within each age-condition combination. These results should be interpreted

---

[3] For those participants that never endorsed retrieving, I cannot rule out that they did not retrieve at all, per se. I can only know that they did not endorse the retrieval strategy.
[4] Including participants that did not retrieve changed the pattern of results, primarily because they were located at the origin, which influenced the results to make them appear stronger than those reported in the text, also resulting in differing Pearson and Spearman correlation coefficients. These correlations are reported in Table 2.

cautiously due to low power. For older adults, retrieval use significantly correlated between the two tasks, $r = .33$, $p = .02$, 95% CI [.06, .56]; $r_s = .31$, $p = .03$, 95% CI [-.13, .39]. This indicates a positive linear and increasing monotonic relationship, where greater retrieval use in the NPLT was slightly indicative of greater retrieval use in the GPLT. Retrieval use between the two tasks did not correlate significantly for young adults, however, $r = -.02$, $p = .91$, 95% CI [-.28, .25]; $r_s = .14$, $p = .29$, 95% CI [-.13, .39]. Retrieval use in the NPLT seemed to have no influence on retrieval use in the GPLT for young adults.

For participants in the overpriced condition, retrieval use between the two tasks correlated significantly, $r =.36$, $p = .01$, 95% CI [.09, .58]; $r_s = .29$, $p = .04$, 95% CI [.02, .53]. I suspect this relationship is driven primarily by older adults in this condition, because their retrieval use was low in both tasks, and the young adults' retrieval was greater than that of the older adults in both tasks. Retrieval use between both tasks for participants in the market-priced condition did not correlate significantly, $r = .19$, $p = .14$, 95% CI [-.07, .44]; $r_s = .21$, $p = .12$, 95% CI [-.06, .45]. Finally, retrieval use did not correlate significantly for any age-condition combination (see Table 2). These results suggest that for young adults, retrieval use was dependent on the task more so than determined by individual retrieval ability per se. The modest relationship between retrieval use for the older adults suggests that individual differences in this sample were predictive of retrieval use.

Post-Task Questionnaires

Analyses will be reported separately by task, conducting two-way ANOVAs comparing age and condition on the post-task questionnaire item. Questions included in the post-task questionnaires are included in the Appendix B. Marginal means are included in Table 2 for both the GPLT and the NPLT.

Global confidence, JOLs, and recall accuracy in the GPLT. Condition differences were found in reported confidence in using memory, or global confidence, where participants in the market-priced condition were more confident than overpriced participants (*M* = 72.43, *SE* = 3.4; *M* = 50.19, *SE* = 3.4, for market-priced and overpriced conditions, respectively), $F(1, 114) = 21.34$, $MS_{Error} = 683.78$, $p < .001$, $\eta^2_G = .16$. Neither age, nor the age by condition interaction were significant, *F*s < 1.

In estimating the likelihood of recalling the price that goes with an item, a judgment of learning (JOL), only condition was significant, where the market-priced group indicated they were more likely to recall the pairs on average compared to the overpriced group (*M* = 81.47, *SE* = 3.16; *M* = 55.41, *SE* = 3.16, for market-priced and overpriced groups, respectively), $F(1, 114) = 33.95$, $MS_{Error} = 589.57$, $p < .0001$, $\eta^2_G = .23$. Again, there was no effect of age nor an age by condition interaction, *F*s < 1.

The market-priced group recalled more items than the overpriced group (*M (SE)* = .79 (.03); *M* (*SE*) = .48 (.03), for the market- and overpriced groups, respectively), $F(1, 114) = 43.40$, $MS_{Error} = 0.069$, $p < .0001$, $\eta^2_G = .28$. Age was not significant, $F(1, 114) = 3.20$, $MS_{Error} = 0.069$, $p = .08$, $\eta^2_G = .03$, nor was the age by condition interaction, *F*(1,

114) = .06. These findings are in line with the pattern of recognition memory performance, where the market-priced group out-performed the overpriced group.

To compare the relative accuracy of the JOLs with recall, Goodman-Kruskal gamma correlations between JOLs and recall (Nelson, 1984) were computed. This measure is an index of the rank order agreement of metacognitive judgments (JOLs) with recall, and can help determine the extent to which a person can effectively discriminate between more or less likely to-be-remembered items (see also Dunlosky, Mueller, & Thiede, 2015; Nelson, 1996). The index computes the correlation between JOL ratings and the binary outcome (correct/incorrect) of recall for each individual, computing the difference of number of discordances from concordances, over the set of rated items. These correlations can then be computed for each between-subjects group and tested for reliable differences. Gammas were reliable (YA-MP: $M = .74$, $SE = .14$; OA-MP: $M = .58$, $SE = .13$; YA-OP: $M = .46$, $SE = .12$; OA-OP: $M = .56$, $SE = .12$), however no condition or age differences emerged, ps > .23, indicating this sample had good metacognitive monitoring of their learned items, but neither condition nor age offered any differences in this. Because the mean gamma for the YA-MP seemed higher than the other groups, post-hoc tests were run to see if any differences emerged. No pairwise comparisons were significant, ps > .42. This finding replicates previous findings using the NPLT, and what I found with the NPLT in the present study (Hertzog & Touron, 2011; Touron & Hertzog, 2004b; Touron et al., 2007).

Global confidence, JOLs, and recall accuracy in the NPLT. To preface, the pattern of results found here replicates prior findings with the NPLT (e.g., Touron &

Hertzog, 2004a, 2004b; Touron et al, 2007). No effects of condition were significant, but the age by condition marginal means are reported in Table 2. Four additional participants were removed from all post-task analyses due to failure to follow instructions (due to comprehension problems with the JOL instructions), in addition to the two participants excluded before due to chance performance in the task. Older adults felt less confident in using the memory strategy than young adults, $M = 84.88$, $SE = 3.14$, $M = 58.49$, $SE = 3.14$, for young and older adults, respectively, $F(1, 110) = 35.21$, $MS_{Error} = 563.2$, $p < .0001$, $\eta^2_G = .24$, which is a consistent pattern with previous work (Touron & Hertzog, 2004a; Touron et al., 2007). Condition was not significant, $F = .24$, nor was the age by condition interaction, $F = 3.38$, $MS_{Error} = 563.2$, $p = .068$, $\eta^2_G = .02$. Although not significant by a post-hoc test, the older adults in the overpriced condition were slightly more confident than those in the market-priced group, ($M = 53.31$, $SE = 4.41$; $M = 63.38$, $SE = 4.49$, for older adults in the market-priced and overpriced conditions, respectively), $t(110) = 1.65$, $p = .36$, $d = 0.44$.

There was a main effect of age in JOLs as well, where YAs had higher confidence in later recall than OAs ($M = 90.18$, $SE = 2.89$; $M = 58.89$, $SE = 2.89$, for YAs and OAs, respectively), $F(1, 110) = 58.22$, $MS_{Error} = 479.13$, $p < .0001$, $\eta^2_G = .35$. Neither condition nor the age by condition interaction was significant, Fs < 1.17. Gamma correlations were calculated. Gammas were reliable for each group (YA-MP: $M = .77$, $SE = .11$; OA-MP: $M = .79$, $SE = .09$; YA-OP: $M = .65$, $SE = .11$; OA-OP: $M = .71$, $SE = .09$), with no differences between them, ps > .77.

Other post-task questionnaire items in the GPLT. In estimating the number of pairs memorized (global accuracy), the market-priced group estimated more memorized items than the overpriced group (*M* = 8.29, *SE* = 0.35; *M* = 5.34, *SE* = 0.35, for the market-priced and overpriced groups, separately), $F(1, 114) = 35.17$, $MS_{Error} = 7.13$, $p < .0001$, $\eta^2_G = .23$. There was no effect of age, $F(1, 114) = 1.05$, $MS_{Error} = 7.13$, $p = .31$, $\eta^2_G = .009$, nor any interaction between age and condition, $F(1, 114) = .001$, $MS_{Error} = 7.13$, $p = .97$. I had expected to find an age difference, based on previous work with the NPLT, but this makes sense given patterns of retrieval and recognition memory performance examined thus far.

Age differences were found when participants estimated how often the memory-only strategy was used, $F(1, 114) = 21.47$, $MS_{Error} = 643.81$, $p < .0001$, $\eta^2_G = .16$. An effect of condition was also found, $F(1, 114) = 38.37$, $MS_{Error} = 643.81$, $p < .0001$, $\eta^2_G = .22$, with no interaction, $F(1, 114) = .01$, $MS_{Error} = 643.81$, $p = .91$. Those in the market-priced group had higher estimations, and young adults had higher estimations (YA-MP: *M* = 77.72, *SE* = 4.7; OA-MP: *M* = 55.53, *SE* = 4.6; YA-OP: *M* = 48.24, *SE* = 4.7; OA-OP: *M* = 27.13, *SE* = 4.6).

No age or condition differences were found in estimated amount of effort in memorizing the grocery-price pairs, age: $F(1, 114) = 0.42$, $MS_{Error} = 783.28$, $p = .52$, $\eta^2_G = .004$; condition: $F(1, 114) = 2.61$, $MS_{Error} = 783.28$, $p = .11$, $\eta^2_G = .022$; age by condition interaction, $F(1, 114)$, $MS_{Error} = 783.28$, $p = .12$, $\eta^2_G = .021$. In estimating the difficulty of the GPLT, condition differences were found, $F(1, 114) = 17.70$, $MS_{Error} = 644.45$, $p < .001$, $\eta^2_G = .13$, where participants in the overpriced condition rated the task as more difficult than those in the market-priced group (YA-MP: $M = 48.97$, $SE = 4.71$; OA-MP: $M = 38.23$, $SE = 4.63$; YA-OP: $M = 66.07$, $SE = 4.71$; OA-OP: $M = 60.47$, $SE = 4.63$). No age differences were found, $F(1, 114) = 3.05$, $MS_{Error} = 644.45$, $p = .08$, $\eta^2_G = .03$, nor was the interaction significant, $F(1, 114) = .3$, $MS_{Error} = 644.45$, $p = .58$.

The trend of older adults seemingly finding the GPLT somewhat easier than the young adults was confirmed in comparing the difficulty of the GPLT to the NPLT (ratings range from 0 to 100, where 0 indicated the GPLT was easier, and 100 indicated the GPLT was harder than the NPLT). Main effects of age, $F(1, 114) = 35.67$, $MS_{Error} = 750.78$, $p < .0001$, $\eta^2_G = .23$, condition, $F(1, 114) = 17.30$, $MS_{Error} = 750.78$, $p < .001$, $\eta^2_G = .10$, and their interaction were significant, $F(1, 114) = 5.42$, $MS_{Error} = 750.78$, $p = .022$, $\eta^2_G = .034$. It seems that young adults in both conditions found the task more difficult than the NPLT, and were not significantly different from one another, $t(114) = 1.28$, $p = .57$, $d = .34$ ($M = 68.38$, $SE = 5.08$; $M = 77.62$, $SE = 5.08$, for market-priced and overpriced conditions, respectively). Older adults exhibited a different pattern. Older adults in the market-priced condition found the task easier than the NPLT, while older adults in the overpriced condition found the task slightly harder but not different from equal difficulties ($M = 26.5$, $SE = 5.00$; $M = 59.23$, $SE = 5.00$, for the market-priced and overpriced conditions, respectively). Post-hoc tests revealed no pairwise differences between YA-MPs, YA-OPS, and OA-OPs, although the YA-OP vs OA-OP comparison was close to significance, $t(114) = 2.58$, $p = .054$, $d = .67$, YA-MP vs YA-OP: $t(114) = 1.28$, $p = .57$, $d = .33$, YA-MP vs OA-OP: $t(114) = 1.28$, $p = .58$, $d = .33$. The OA-MP group was significantly lower compared to all other groups, ts > 4.6, ps < .001, ds > 1.19.

This pattern of results is exciting, as I expected older adults to find the market-priced condition of the GPLT easier than the NPLT, which the older adults in this condition confirmed. This could suggest that schematic support was more beneficial for older adults relative to young adults, and that awareness of the comparative difficulty influenced strategic behavior in the GPLT, particularly for older adults in the market-priced condition.

There were no age or condition differences in familiarity with the grocery items used in the task, *F*s <= 1.28, MS_{Error} = 601.65, *p*s > .25, η^2_G <= .011. Older adults indicated that they went shopping more often than young adults, $F(1, 114) = 6.79$, MS_{Error} = 760.79, $p = .01$, $\eta^2_G = .06$, $M = 65.29$, $SE = 3.62$, $M = 78.53$, $SE = 3.56$, for young and older adults, respectively. This replicates Castel's (2005) finding that older adults went grocery shopping more often than young adults. While curious as a peek into the participants' lifestyles, this difference doesn't have any predictive value with regard to other findings (to preview), as Castel (2005) also found.

In estimating how well the participants felt they did overall, an interaction between age and condition was significant, $F(1, 114) = 4.49$, $MS_{Error} = 408.25$, $p = .036$, $\eta^2_G = .037$, although age, $F(1, 114) = .07$, MS_{Error} = 408.25 $p = .79$, and condition were not significant, $F(1, 114) = 2.09$, MS_{Error} = 408.25 $p = .15$, $\eta^2_G = .02$. Young adults in the market-priced condition reported numerically, but not significantly higher estimates of their performance (YA-MP: $M = 75.38$, $SE = 3.75$; OA-MP: $M = 66.5$, $SE = 3.68$; YA-OP: $M = 62.10$, $SE = 3.75$; OA-OP: $M = 69.00$, $SE = 3.68$). The difference between the two young adult conditions (YA-MP vs YA-OP) approached significance, $t(114) = 2.50$,

p = .07, d = .66; all other comparisons were not significantly different, ps > .33, ds < .44. Participants in the market-priced condition were more satisfied with their performance than those in the overpriced condition (M = 71.3, SE = 3.28; M = 61.43, SE = 3.28, for market-priced and overpriced conditions, respectively). Accordingly, there was a significant effect of condition in reporting satisfaction of their performance in the experiment, $F(1, 114) = 4.49$, $MS_{Error} = 637.21$, $p = .036$, $\eta^2_G = .037$; age and the interaction were not significant, Fs < 1.89, $MS_{Error} = 637.21$, ps > .17, η^2_G s < .02.

No significant differences emerged when participants estimated if they could have done better if they had tried harder, Fs <= 3.8, $MS_{Error} = 1140$, ps > .05, $\eta^2_G < .03$. Older adults found the experiment less fatiguing than young adults, $F(1, 114) = 28.13$, MS_{Error} = 835.86, $p < .0001$, $\eta^2_G = .19$ (YA-MP: M = 70.14, SE = 5.37; OA-MP: M = 39.67, SE = 5.38; YA-OP: M = 67.00, SE = 5.37; OA-OP: M = 41.00, SE = 5.28). Interestingly, OAs on average took three hours to complete the experiment, while YAs were able to complete the study in under two hours. However, this finding is typical in NPLT studies (although not reported in published reports). There were no age or condition differences in feelings of stress or tension, Fs <= 1.06, $MS_{Error} = 737.58$, ps > .31, $\eta^2_G < .009$.

Other post-task questionnaire items in the NPLT. In estimating the total number of pairs they had memorized, there was a main effect of age, where YAs estimated more pairs memorized than OAs (M = 9.95, SE = 0.338; M = 6.56, SE = 0.338, for YAs and OAs, respectively), $F(1, 110) = 49.89$, $MS_{Error} = 6.54$, $p < .0001$, $\eta^2_G = .31$. Older adults estimated using the memory-only strategy less often than young adults, $F(1, 110) = 87.43$, $MS_{Error} = 690.34$, $p < .0001$, $\eta^2_G = .44$; M = 78.71, SE = 3.48, M = 32.68,

$SE = 3.48$, for young and older adults, respectively. No age differences were found in reported effort in memorizing the noun pairs, $F(1, 110) = 1.33$, $MS_{Error} = 764.15$ $p = .25$, $\eta^2_G = .012$. Finally, an age difference was found in estimating the difficulty of the NPLT, where Both age groups found the task relatively easy, although the older adults found the task somewhat harder ($M = 24.8$, $SE = 3.6$; $M = 37.5$, $SE = 3.6$, for young and older adults, respectively), $F(1, 110) = 6.21$, $MS_{Error} = 738.97$, $p = .01$, $\eta^2_G = .05$. Neither condition effects nor age by condition interactions were significant for these analyses, Fs $< .5$, $MS_{Error} = 738.97$, ps $> .5$.

Correlations between GPLT retrieval use and metacognitive measures. A suite of motivated correlations was run with retrieval use in the final block of the GPLT, which are reported in Table 4. Correlations were run separately by age group, consistent with previous work (Touron & Hertzog, 2004b; Touron et al. 2007); both sets of correlations will be reported together, beginning with post-task questionnaire items. Alpha was corrected to .002 (.05/24 correlations). I expected to find positive correlations with end retrieval use for older adults in estimated number of memorized items, global confidence (memory strategy confidence), JOLs, cued recall and recognition memory accuracy, with no predictions regarding correlations with PBMI items, given the small correlation usually reported (~ $r = .15$, Beaudoin & Desrichard, 2011).

Number of estimated memorized pairs correlated significantly with end retrieval for both young, $r = .47$, 95% CI [.22, .65], $p < .001$, and older adults, $r = .59$, 95% CI [.40, .74], $p < .0001$. Confidence in using the memory strategy (global confidence) also correlated positively for young, $r = .47$, 95% CI [.22, .66], $p < .001$ and older adults, $r =$

.57, 95% CI [.37, .72], $p < .0001$. Estimation of the proportion of time the retrieval strategy was used was also significant for young, $r = .42$, 95% CI [.17, .63], $p < .002$, and older adults, r = .71, 95% CI [.56, .82], $p < .0001$. No significant correlations emerged between end retrieval use and the PBMI items, consistent with previous work with the NPLT (Touron & Hertzog, 2004a).

JOLs significantly correlated with end retrieval use for both young, $r = .57$, 95% CI [.35, .73], $p < .001$, and older adults, $r = .67$, 95% CI [.48, .79], $p < .001$. Cued recall was also significantly correlated with retrieval use for young, $r = .54$, 95% CI [.32, .71], $p < .001$, and older adults, $r = .68$, 95% CI [.49, .81], $p < .001$. End retrieval use also significantly correlated with *d’* for both young, $r = .57$, 95% CI [.37, .72], $p < .001$, and older adults, r = .62, 95% CI [.43, .75], $p < .001$.

Correlations between NPLT Retrieval use and metacognitive measures. Correlations between retrieval use at the end of the NPLT was correlated with post-task questionnaire items and (meta-)memory measures, separately by age, and can be found in Table 3. The pattern of predictions was the same as that for the GPLT, with positive correlations expected for older adult retrieval use with estimated number of memorized items, global confidence, JOLs, recall and recognition memory, and no predictions for the PBMI. Alpha was corrected to .003125 (.05/16 correlations).

Estimated number of memorized pairs correlated significantly with end retrieval use in the NPLT for both young, $r = .52$, 95% CI [.29, .69], $p < .001$, and older adults, $r = .50$, 95% CI [.27, .68], $p < .001$. Older adults’ confidence in using the memory strategy was not significant by this strict alpha, $r = .36$, 95% CI [.11, .58], $p = .007$, however it is in the range of previously reported correlation values in noun-pair studies, *r*s ~= .35 (Touron & Hertzog, 2004a, 2009; Touron et al., 2007). Young adults' confidence in using the memory strategy was not significant by my alpha cut-off either, but would meet traditional significance, $r = .31$, 95% CI [.06, .53], $p = .02$., which is somewhat surprising considering the typical strength of this correlation in young adults is often close to zero (e.g., Touron et al, 2007). Estimation of how often the memory strategy was used correlated significantly with retrieval use for older adults, $r = .61$, 95% CI [.40, .75], $p < .001$, but not significantly with young adults, $r = .36$, 95% CI [.11, .56], $p = .006$. No items from the PBMI significantly correlated with end retrieval use.

JOLs significantly correlated with end retrieval use for young adults, $r = .49$, 95% CI [.25, .67], $p < .001$, but, unexpectedly, not for older adults, $r = .37$, 95% CI [.05, .61], p = .02. Cued-recall accuracy did significantly correlate with both young, $r = .42$, 95% CI [.17, .62], $p < .001$, and older adults, $r = .55$, 95% CI [.28, .74], $p < .001$. End retrieval use significantly correlated with *d'* for both young, $r = .79$, 95% CI [.66, .87], $p < .001$, and older adults, $r = .68$, 95% CI [.50, .79], $p < .001$.

Age differences in the PBMI. Age differences only emerged for two items: change in memory over the past 10 years (retrospective change), $t(103.18) = 4.1$, $p <.001$, $d = 1.66$, and expected change in memory over the next 10 years (prospective change), $t(101.94) = 6.04$, $p < .0001$, $d = 1.14$. Comparisons of global memory ability, relative standing (memory compared adults of all ages, and to adults of their age), control over memory, prospective control, future control over memory, and specific memory abilities (i.e., memory for people, faces, names, etc.) failed to show age differences, all *p*s > .46.

Failure to find age-differences has been reported previously (Hines et al., 2014, Touron & Hertzog, 2004a), but they have been found as well (Lineweaver & Hertzog, 1998). Post-task questionnaire results were correlated with the PBMI results, which are reported in Appendix C. The correlations were not significant but for a few exceptions, due to a strict alpha that was adopted because of the exploratory nature of these analyses. Correlations of MSE with memory tests did not meet significance.

CHAPTER IV

DISCUSSION

The present study demonstrated that older adults' strategic use of memory in a cognitive skill learning task can be adaptively altered depending on the nature of the task materials. The typical age-related dissociations in memory-based strategy use were found in the NPLT, where young adults readily adopted the memory-based strategy, and older adults were slow to adopt this strategy. However, this pattern was not found in the GPLT. Young and older adults in the GPLT readily shifted to a memory-based strategy for the market-priced materials, which supplied schematic support in that a plausible association between the groceries and prices existed prior to the participants entering the lab. For the overpriced materials, young and older adults adopted the memory-based strategy later in the task and to a lesser extent than what was observed for the market-priced materials.

This study indicates that older adults are not reluctant to adopt memory-based strategies across all situations (e.g., Touron, 2015). These findings suggest that older adults' confidence in using a memory-based strategy is critical to whether it is used, which will vary depending on the material to be learned. Older adults have shown no differences from young adults on associative memory measures when the pairs are related (Old & Naveh-Benjamin, 2008), in part due to the ease with which one can incorporate the pairs into existing schematic frameworks. This ease translated into older adults' increased confidence in using their memory and in increased use of memory in the GPLT.

Surprisingly, young adults in the overpriced condition of the GPLT showed far less retrieval use than would be expected from previous work (e.g., Touron et al., 2001, 2004; Touron & Hertzog, 2004a, 2004b, 2009; Touron et al., 2007, but see Ackerman & Woltz, 1994), and accordingly I predicted that retrieval use would not differ between the young adult groups in either the market-priced or overpriced conditions. Additionally, the prices used in the market-priced condition were accurate to their real-life groceries, but the selection of groceries was constrained such that the prices could reasonably apply to any of the groceries. I did not expect that participants would know the specific prices for the groceries, but rather would feel that the prices were reasonable within expectation. In the overpriced condition, the prices were expected to be unreasonable for all groceries involved. It could be that the unreasonableness of the overpriced groceries interfered with learning, which could be why these young adults expressed lower confidence that they could successfully use the memory strategy.

The overpriced groceries could have caused interference through means of processing the items. Because the items and prices were drawn from the same category, organizational processing would be ineffective (e.g., Hunt & Lamb, 1991), and successful discrimination of pairs would be reliant on processing the pairs individually. Individual-pair processing in the overpriced condition would be inferior to individual-pair processing in the market-priced condition, given that the pairs could be incorporated into an existing associative network in the market-priced condition. That is, participants are likely better at implementing a distinctive processing strategy in the market-priced condition, where similarities and differences between pairs could potentially be meaningful, whereas in the overpriced condition, judging similarities and differences between pairings is more difficult.

Ackerman and Woltz (1994) demonstrated similar findings when comparing learning rates in the NPLT for unrelated noun-pairs when there was no relation between any pairs at all, and when the pairs were drawn from two categories (e.g., Volcano-Doll, Lawn-Planet vs. Doctor-Rug, Nurse-Chair). When there was no relation between any of the pairs, learning occurred much faster than when the pairs shared some associative relationship. The list-wide associations (despite individual pairs being unrelated) seems detrimental to young adults' learning rates. For older adults in the context of the item-pair lookup tasks (NPLT and GPLT), unrelated pairs, regardless of list-wide relatedness, will hurt learning rates when compared to learning related pairs. If I implemented a between-measures experiment where young and older adults learned unrelated noun-pairs all drawn from different categories (as done here) and from only two different categories (as in Ackerman & Woltz, 1994), I would expect to see young adults show slower learning in the two-category pair condition than an all different category condition. Older adults would presumably show worse learning than young adults in both conditions.

A surprising finding was the modest correlation between retrieval use in the final blocks of both the GPLT and NPLT ($r = .29$). Several older adults did not retrieve at all in either task (4 participants), and several went on to retrieve in the GPLT after not retrieving in the NPLT (7 participants). In sum, 11 participants never endorsed retrieval in the NPLT (~37% of the sample), which is similar to a report from Touron (2015), finding around a third of the sample not retrieving at all in an unpublished NPLT study aimed at finding individual differences in strategy use. The findings from the current study indicate that like with associative memory, not all kinds of cognitive skill learning are the same for older adults. The retrieval rate for participants who were retrieving modestly increased from what was measured for the NPLT, and several of those who did not retrieve in the NPLT retrieved in the GPLT (although I cannot offer any explanation as to why this change occurred for these participants).

These data indicate that participants' retrieval usage is sensitive to the material being learned, where schematic support can change participants' strategic approach in two item-pair lookup tasks. This suggests that young and older adults' strategic sets are flexible and task-dependent. Pertinent to this, Rogers and Gilbert (1997) showed that prior practice influenced the number of older adults retrieving in the NPLT. In the present study it is impossible to separate the effects of practice on retrieval use in the GPLT, and so I may have found greater numbers of older adults retrieving than would have been found without prior practice. It is certain that practice does not have uniform carryover effects on retrieval use: older adults in the market-priced condition, but not the overpriced condition increased retrieval use; young adults' retrieval use decreased in the overpriced condition and retained roughly the same pattern of retrieval use in the market-priced condition when compared with the NPLT. If these results were solely a result of practice, I would have expected to see gross increases in retrieval use in the GPLT by all participants, given that it is the most efficient strategy. Ruling out practice effects will require only testing participants on the GPLT.

It is harder to explain why recognition memory performance differed between age groups in the NPLT, but not in the GPLT, where instead there were differences between conditions (market-priced vs. overpriced). Finding age differences in recognition memory after completing the NPLT is consistent with prior work (Touron & Hertzog, 2004b; Touron et al., 2007). Condition differences have not been found before, whether manipulating the memory and scanning loads (Touron & Hertzog, 2004b) or given incentives to retrieve (Touron et al., 2007), indicating that differences in older adults' retrieval use during the task did not translate into differences in their noun-pair learning measured by the recognition memory test, but prior work has only used the NPLT and so it was not clear how the GPLT would impact results.

Given these past results, it would not have been surprising to find age differences in the overpriced recognition memory test, where the young adults outperform the older adults. Inspecting the differences in recognition memory between the NPLT (Figure 6) and the GPLT (Figure 2), young adults' recognition memory in the overpriced condition was lowered (M (s) = .91 (.09); M (s) = .82 (.14), for the NPLT and GPLT, respectively) whereas older adults' recognition memory remained roughly the same (M (s) = .78 (.15); M (s) = .78 (.15), for the NPLT and GPLT, respectively). This finding would need to be replicated before any firm conclusions can be drawn, considering that sampling error, practice effects, and schematic support could all be contributing to this effect. Follow-up research would need to clarify the relationship between recognition memory performance and retrieval use in the task, possibly by varying the types of material to be learned, with varying levels of schematic support.

A note on automaticity is warranted. Logan (1988) interpreted automaticity as memory-based responding, which could occur as a broad task-set (i.e., memory-based responding for every item), and just for specific items (i.e., memory-based responding for some, but not all items). In this sense, almost every participant achieved some form of automaticity in the current study (except for those who never retrieved in either task). The schematic support manipulation worked as intended for older adults in the market-priced condition of the GPLT, where they reached greater levels of memory-based responding than those in the overpriced condition. The level of memory-based responding for the older adults in the market-priced condition was statistically equivalent with young adults in the market-priced condition by the end of practice, and so in this sense I can say that older adults demonstrated preserved automaticity (cf. Rawson & Touron, 2009, 2015).

There is evidence to suggest that memory-based responding and automaticity can be distinctly identified in a single experimental session (Tenison & Anderson, 2015). Using hidden Markov-models of response time data, Tenison and Anderson (2015) were able to identify three distinct states of responding in a novel arithmetic task, favoring evidence for a three-phase model (i.e., distinct phases corresponding to algorithmic, memory-based, and automatic responding) over a two-phase model that only distinguishes between algorithmic and memory-based responding. Although the concern of the present experiment was with the shift from algorithmic to memory-based responding, it would be simple to adapt the current tasks to investigate this other kind of strategy shift – slightly more than two times the current number of trials in a single task would need to be added.

The pattern of results is inconsistent with prominent skill acquisition theories in several senses. Diverging onsets of shifts-to-retrieval in young adults were found in the GPLT, depending on the item-pair type. According to CMPL (Rickard, 1997), the onset of memory-based responding should be approximately equal regardless of algorithm difficulty within a task (e.g., artificial arithmetic problems with differing sizes of addends should shift to retrieval at approximately the same time despite larger addends being more difficult). The scanning algorithm in the GPLT tasks should be approximately equal in difficulty (although it could be plausible that scanning takes slightly longer in the overpriced condition as there are more characters to scan), and so CMPL would predict the onset to retrieval occurring at the same time.

Instance theory is not equipped to handle these data either, because it cannot account for volitional strategy selection (Logan, 1988). The fundamental assumption of instance theory is that the algorithm and retrieval strategies proceed in parallel, where the number of stored instances determines the winning strategy. It’s not clear how instance theory would be able to account for these data without abandoning this assumption; to do so would require radical reconfiguration of the theory.

Although CMPL does allow for strategy selection, beyond the strengthening effects of practice, no other selection mechanism is specified. To date, the retrieval reluctance hypothesis has been successful in explaining older adults’ lower-than-expected usage of the retrieval strategy, indicating that volitional choice can dictate strategy usage, rather than being solely determined by bottom-up learning mechanisms like instance theory would predict (e.g., Touron, 2015). The results of the present study indicate that volitional strategy choice can be an important factor in determining strategic behavior in skill learning tasks, but it may be something of an overriding presence that operates when one's confidence in the retrieval strategy is low. This may be due in part to the difficulty of learning the pair materials. Understanding situations when a volitional 'override' may affect strategic choice will be important for future research to explore.

These data suggest that such a volitional 'override' is not always a factor in young adults' strategic choice in skill learning tasks, and so existing theories may do well in describing their performance in such cases, although they do not perfectly apply. CMPL has been particularly effective in capturing young adults' performance in skill learning tasks, although it is at a loss to explain (for example) why a participant would not shift at all (as seen in experiment 1 in Rickard, 1997) or why certain classes of items in a task would shift to retrieval sooner than others, as shown in Rickard's (1999) reanalysis of Palmeri's (1997) numerosity judgment task.

To account for this lack of strategy-selection mechanism, Rickard and colleagues (1997; Bajic & Rickard, 2009, 2011) have suggested that a feeling-of-knowing (FOK) judgment may guide strategy choice, as the Source Activation Confusion model suggests (Schunn et al., 1997; Reder & Ritter, 1992). However, Hertzog and Touron (2011) demonstrated that FOK judgments are not entirely commensurate with strategy choices, even for young adults. By having participants either make a high or low FOK judgment or choose which strategy to use (scanning or retrieval) before a NPLT trial (between-subjects), Hertzog and Touron found diverging patterns of strategy usage between young and older adults (see their Figures 3 and 6). If an FOK judgment and pre-trial strategy choice were commensurate, they would have expected to see approximately equal proportions of trials where young and older adults made high FOKs and chose to retrieve; observing approximately equal proportions of high FOK trials and pre-trial retrieval choices would have indicated that older adults' lower proportions of retrieval in the NPLT is a function of learning, not reluctance to use the retrieval strategy.

They did not observe this pattern. Older adults made higher proportions of high FOK judgments than older adults choosing to use the retrieval strategy before a trial. Young adults' high FOK judgments and pre-trial retrieval-strategy choices matched almost perfectly for intact trials, where the cued-pair matched an existing pair in the lookup table. However, the proportion of young adults' high FOK judgments was considerably lower than pre-trial retrieval choices for rearranged pairs, which could signal differential reactivity to the experimental conditions. On the other hand, it could support the idea that volitional strategy choice can selectively applied, especially in situations where a participant experiences uncertainty regarding the effectiveness of the retrieval strategy. This idea is also supported by the fact that later endorsement of the retrieval strategy by participants in the FOK condition following a high FOK judgment was lower for rearranged trials than intact trials. This is to say that FOK judgments are a plausible mechanism in making strategy choices, and very well could be made with each strategy choice, however, they do not map one-to-one with each strategy choice.

Limitations of this study must be noted. I cannot rule out item effects contributing to memorability from this experiment – only one pairing was used for the noun-pairs and for the respective conditions of the grocery task. Item effects in the market-priced condition may have allowed for greater elaboration when participants viewed the grocery-price pairs. Several participants informally noted that certain pairings seemed discrepant with their personal schemas (e.g., thinking that the price of soda was too high). This discrepancy could have led to certain items having greater distinctiveness than was anticipated, making learning easier for some pairs in the market-priced condition than others. Follow-ups should counterbalance pairings of stimuli to counteract item effects. Market-priced items would still be expected to have mnemonic superiority over overpriced items, but specific item effects can be eliminated.

Furthermore, I cannot rule out order effects from first completing the NPLT in determining performance in the GPLT. Participants must only complete the GPLT in order to determine that schematic support, and not prior practice, is influencing performance. Additionally, I cannot guarantee that the same effects of price-type on retrieval use would be observed if participants completed a version of the GPLT with both market- and overpriced grocery items; participants may be able to more easily learn the overpriced pairs if they can distinguish them from the market-priced pairs. It is also uncertain if different kinds of item-pair types would produce the same results. Grocery-price pairs might be particularly difficult to learn, especially so if the pairings are as inconceivable as the overpriced pairs. Different materials must be tested to determine to determine the generalizability of this effect.

That the young adults in the overpriced condition exhibited apparent retrieval reluctance suggests that deficits in older adults' associative learning ability or metacognitive monitoring may have been overstated in previous studies. It's undeniable that an associative deficit is largely influential in older adults' strategy choice, and it has been shown that older adults can have errant beliefs about the efficiency of the retrieval strategy (Hertzog, Touron, & Hines, 2007; Herzog & Touron, 2011) – the present work calls into question just what is a 'deficit' in older adults' strategic abilities and what is a general inefficiency in human strategic behavior. Clarifying this issue will have value towards understanding how our cognition changes with aging and provide better insight into aiding older adults' learning.

This study demonstrated strategic choice in cognitive skill learning is influenced by schematic support, for both young and older adults. It also identified that confidence in memory-based responding can also matter for young adults, rather than just for older adults, resulting in different levels of memory-based responding according to the level of schematic support provided. Identifying when and why young and older adults' performance in cognitive skill learning tasks differs and does not differ would be valuable towards a better understanding of learning and how it changes across the lifespan.

REFERENCES

Ackerman, P. L., & Woltz, D. J. (1994). Determinants of learning and performance in an associative memory/substitution task: Task constraints, individual differences, volition, and motivation. *Journal of Educational Psychology, 86*, 487–515.

Adams, C., Smith, M. C., Nyquist, L., & Perlmutter, M. (1997). Adult age-group differences in recall for the literal and interpretive meanings of narrative text. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 52*(4), P187-P195.

Bajic, D., & Rickard, T. C. (2009). The temporal dynamics of strategy execution in cognitive skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35(*1), 113.

Bajic, D., & Rickard, T. C. (2011). Toward a generalized theory of the shift to retrieval in cognitive skill learning. *Memory & cognition, 39*(7), 1147-1161.

Bakeman, R. (2005). Recommend effect size statistics for repeated measure designs. *Behavior Research Methods*, *37*(3), 379–384.

Bandura, A. (1989). Human agency in social cognitive theory. *American psychologist*, *44*(9), 1175.

Beaudoin, M., & Desrichard, O. (2011). Are memory self-efficacy and memory performance related? A meta-analysis. *Psychological Bulletin*, *137*(2), 211–241. http://doi.org/10.1037/a0022106

Beaudoin, M., & Desrichard, O. (2017). Memory self-efficacy and memory performance in older adults: The mediating role of task persistence. *Swiss Journal Of Psychology*, *76*(1), 23-33. doi:10.1024/1421-0185/a000188

Berry, Jane M. "Memory Self-Efficacy in Its Social Cognitive Context." In *Social Cognition and Aging*, edited by Thomas M. Hess and Fredda Blanchard-Fields, 69-96. San Diego: Academic Press, 1999.

Bottiroli, S., Dunlosky, J., Guerini, K., Cavallini, E., & Hertzog, C. (2010). Does task affordance moderate age-related deficits in strategy production? *Aging, Neuropsychology, and Cognition, 17*(5), 591-602.

Castel, A. D. (2005). Memory for grocery prices in younger and older adults: the role of schematic support. *Psychology and Aging*, *20*(4), 718-721. http://doi.org/10.1037/0882-7974.20.4.718

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition, 24*, 403–416.

Cohen, J. (1992). Quantitative Methods in Psychology. *Psychological Bulletin*, *112*(1), 155–159. http://doi.org/10.1037/0033-2909.112.1.15

Cohn, M., Emrich, S. M., & Moscovitch, M. (2008). Age-Related Deficits in Associative Memory: The Influence of Impaired Strategic Retrieval. *Psychology and Aging*, *23*(1), 93–103. http://doi.org/10.1037/0882-7974.23.1.93

Connor, L. T., Dunlosky, J. D., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, *12*, 50-71.

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transaction of the Royal Society of London*, Series B, 302, 341-359.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities, mechanisms and performance* (pp. 409-422). Amsterdam: North-Holland and Elsevier.

Craik, F. I. M., & Bosman, E. A. (1992). Age-related changes in memory and learning. In H. Bouma & J. A. M. Graafman (Eds.), *Gerontechnology* (pp. 79-92). Amsterdam: IOS Press.

Craik, F. I. M., & Jennings, J. M. (1992). Human Memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (p. 51-110). Hillsdale, NJ: Erlbaum.

Delacre, M., Lakens, D., & Leys, C. (2017, February 20). Why Psychologists Should by Default Use Welch's t-test Instead of Student's t-test (in press for the International Review of Social Psychology). Retrieved from osf.io/sbp6k

Dunlosky, J., & Hertzog, C. (1998). Aging and deficits in associative memory: What is the role of strategy production? *Psychology and Aging, 13*, 597–607.

Dunlosky, J., & Hertzog, C. (2001). Measuring strategy production during associative learning: The relative utility of concurrent versus retrospective reports. *Memory & Cognition, 29*, 247–253.

Dunlosky, J., Hertzog, C., & Powell-Moman, A. (2005). The contribution of mediator-based deficiencies to age differences in associative learning. *Developmental Psychology*, *41*(2), 389.

Dunlosky, J., Kubat-Silman, A. K., & Hertzog, C. (2003). Training monitoring skills improves older adults' self-paced associative learning. *Psychology and Aging*, *18*(2), 340.

Dunlosky, J., Mueller, M. L., & Thiede, K. W. (2015). Methodology for Investigating Human Metamemory. In J. Dunlosky & S. (Uma) K. Tauber (Eds.), *The Oxford Handbook of Metamemory* (pp. 1–28). Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199336746.013.14

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211–245.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Frank, D. J., Jordano, M. L., Brown, K., & Touron, D. R. (2016). Older adults' use of retrieval strategies in everyday life. *Gerontology*, 4–8. http://doi.org/10.1159/000446277

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology, 44*, 339–387.

Hertzog, C. (2015). Aging and Metacognitive Control. In J. Dunlosky & S. (Uma) K. Tauber (Eds.), *The Oxford Handbook of Metamemory* (pp. 1–41). Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199336746.013.31

Hertzog, C., & Hultsch, D. F. (2000). Metacognition in adulthood and old age. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 417–466). Mahwah, NJ: Erlbaum.

Hertzog, C., Lineweaver, T. T., & Hines, J. C. (2014). Computerized assessment of age differences in memory beliefs. *Perceptual and Motor Skills*, *119*(2), 609–628. http://doi.org/10.2466/03.10.PMS.119c23z4

Hertzog, C., McGuire, C. L., Horhota, M., & Jopp, D. (2010). Does believing in "use it or lose it" relate to self-rated memory control, strategy use, and recall? *The International Journal of Aging and Human Development*, *70*(1), 61-87.

Hertzog, C., McGuire, C. L., & Lineweaver, T. T. (1998). Aging, attributions, perceived control, and strategy use in a free recall task. *Aging, Neuropsychology, and Cognition*, *5*(2), 85-106.

Hertzog, C., Touron, D. R., & Hines, J. C. (2007). Does a time-monitoring deficit influence older adults' delayed retrieval shift during skill acquisition? *Psychology and aging*, *22*(3), 607.

Hertzog, C., & Touron, D. R. (2011). Age differences in memory retrieval shift: Governed by feeling-of-knowing? *Psychology and Aging*, *26*(3), 647–660. http://doi.org/10.1037/a0021875

Hess, T. M., & Emery, L. (2012). Memory in context. *Memory and Aging: Current Issues and Future Directions*, 183.

Hess, T. M., Rosenberg, D. C., & Waters, S. J. (2001). Motivation and representational processes in adulthood: the effects of social accountability and information relevance. *Psychology and Aging, 16*(4), 629.

Hess, T. M., & Tate, C. S. (1991). Adult age differences in explanations and memory for behavioral information. *Psychology and Aging, 6*(1), 86.

Hines, J., Hertzog, C., & Touron, D. (2012). A prelearning manipulation falsifies a pure associational deficit account of retrieval shift during skill acquisition. *Aging, Neuropsychology, and Cognition, 19*, 449–478.

Hoyer, W. J., Cerella, J., & Onyper, S. V. (2003). Item learning in cognitive skill training: Effects of item difficulty. *Memory & Cognition, 31*(8), 1260–1270.

Hülür, G., Hertzog, C., Pearman, A. M., & Gerstorf, D. (2015). Correlates and moderators of change in subjective memory and memory performance: Findings from the health and retirement study. *Gerontology*, *61*(3), 232–240. http://doi.org/10.1159/000369010

Hunt, R. R., Lamb, C. A. (2001). What causes the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1359-1366.

Kausler, D. H. (1994). *Learning and memory in normal aging*. San Diego, CA: Academic Press.

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, *135*(1), 36–69. http://doi.org/10.1037/0096-3445.135.1.36

Kuhlmann, B. G., & Touron, D. R. (2012). Mediator-based encoding strategies in source monitoring in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(5), 1352.

Kuhlmann, B. G., & Touron, D. R. (2016). Aging and memory improvement through semantic clustering: The role of list-presentation format. *Psychology and Aging, 31*(7), 771.

Lachman, M. E., & Andreoletti, C. (2006). Strategy use mediates the relationship between control beliefs and memory performance for middle-aged and older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *61*(2), 88-94.

Lachman, M. E., Andreoletti, C., & Pearman, A. (2006). Memory control beliefs: How are they related to age, strategy use and memory improvement? *Social Cognition*, *24*(3), 359-385.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*, *4*, 863.

Lamson, N., and Rogers, W. A. (2008). Assessing age-related patterns in strategy selection on a mathematical problem-solving task. *Journal of Gerontology: Psychological Science*, 63, 146–155.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational behavior and human performance*, *20*(2), 159-183.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual review of psychology*, *42*(1), 333-376.

Lineweaver, T. T., & Hertzog, C. (1998). Adults' efficacy and control beliefs regarding memory and aging: Separating general from personal beliefs. *Aging, Neuropsychology, & Cognition, 5*, 264–296.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Murphy, M. D., Sanders, R. E., Gabriesheski, A. S., & Schmitt, F. A. (1981). Metamemory in the aged. *Journal of Gerontology, 36*, 185–193.

Murphy, M. D., Schmitt, F. A., Caruso, M. J., & Sanders, R. E. (1987). Metamemory in older adults: The role of monitoring in serial recall. *Psychology and Aging, 2*, 331–339.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1170–1187.

Naveh-Benjamin, M., Craik, F. I. M., & Ben-Shaul, L. (2002). Age-related differences in cued recall: Effects of support at encoding and retrieval. *Aging, Neuropsychology, and Cognition, 9*(4). http://doi.org/10.1076/anec.9.4.276.8773

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological bulletin*, *95*(1), 109.

Nelson, T. O. (1996). Gamma is a measure of the accuracy of predicting performance on one item relative to another item, not of the absolute performance on an individual item: Comments on Schraw. *Applied Cognitive Psychology*, *10*(3), 257-260.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125–141). San Diego, CA: Academic Press.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(2), 324.

Onyper, S. V, Hoyer, W. J., & Cerella, J. (2008). Effects of item difficulty on the retrieval of solutions during cognitive skill acquisition: age differences. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *15*(3), 358–83.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological methods*, *8*(4), 434.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, *60*(4), 437-447.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org/.

Rahhal, T. A., & Hasher, L. (1998). It's not all downhill: Age differences in explicit memory performance can be eliminated by instructions. In *Cognitive Aging Conference, Atlanta, GA.*

Rawson, K. A., & Touron, D. R. (2009). Age differences and similarities in the shift from computation to retrieval during reading comprehension. *Psychology and Aging, 24*, 423–437. http://dx.doi.org/10.1037/ a0016044

Rawson, K. A., & Touron, D. R. (2015). Preservation of Memory-Based Automaticity in Reading for Older Adults. *Psychology and Aging, 30*(4)*,* 809.

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, memory, and cognition*, *18*(3), 435.

Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General, 126*, 288–310.

Rickard, T. C. (1999). A CMPL alternative account of practice effects in numerosity judgment tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(*2*), 532–542.

Rogers, W. A. & Gilbert, D. K. (1997). Do performance strategies mediate age-related differences in associative learning? *Psychology and Aging, 12*(4), 620-633.

Rogers, W. A., Hertzog, C., & Fisk, A. D. (2000) An individual differences analysis of ability and strategy influences: Age-related differences in associative learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 359–394.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. http://doi.org/10.1037/0033-295X.103.3.403

Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty* (2nd ed.). Mahwah, NJ: Erlbaum.

Schunn, C. D., Reder, L. M., Nhouyvanisvong, A., Richards, D. R., & Stroffolino, P. J. (1997). To calculate or not to calculate: A source activation confusion model of problem familiarity's role in strategy selection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 3–29.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review*, *84*(2), 127.

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2017). afex: Analysis of Factorial Experiments. R package version 0.18-0. https://cran.r-project.org/package=afex

Tenison, C., & Anderson, J. R. (2016). Modeling the distinct phases of skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(5), 749.

Touron, D. R. (2015). Memory avoidance by older adults: When "old dogs" won't perform their "new tricks." *Current Directions in Psychological Science*, *24*(3), 170–176. http://doi.org/10.1177/0963721414563730

Touron, D. R., & Hertzog, C. (2004a). Distinguishing age differences in knowledge, strategy use, and confidence during strategic skill acquisition. *Psychology and Aging, 19*, 452–466.

Touron, D. R., & Hertzog, C. (2004b). Strategy shift affordance and strategy choice in young and older adults. *Memory & Cognition, 32*, 298–310.

Touron, D. R., & Hertzog, C. (2009). Age differences in strategic behavior during a computation-based skill acquisition task. *Psychology and Aging, 24*, 574–585.

Touron, D. R., Swaim, E. T., & Hertzog, C. (2007). Moderation of older adults' retrieval reluctance through task instructions and monetary incentives. *Journal of Gerontology: Psychological Sciences*, *62B*(3), 149–155.

Umanath, S., & Marsh, E. J. (2014). Understanding How Prior Knowledge Influences Memory in Older Adults. *Perspectives on Psychological Science*, *9*(4), 408–426. http://doi.org/10.1177/1745691614535933

Van Overschelde, J. P., Rawson, K. A., Dunlosky, J., & Hunt, R. R. (2005). Distinctive processing underlies skilled memory. *Psychological Science*, *16*(5), 358–361. http://doi.org/10.1111/j.0956-7976.2005.01540.x

Zachary, R. A., & Shipley, W. C. (1986). *Shipley institute of living scale: Revised manual*. WPS, Western Psychological Services.

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 293–358). Mahwah, NJ: Erlbaum

APPENDIX A

TABLES AND FIGURES

Table 1

Demographic Information

Measure	Young MP	Young OP	Old MP	Old OP
Age (years)	19.33 (0.32) [a]	19.53 (0.31) [a]	67.89 (0.75) [b]	68.17 (0.76) [b]
Education	12.80 (0.33) [a]	12.86 (0.33) [a]	16.24 (0.34) [b]	16.20 (0.33) [b]
No. medications	0.56 (0.25) [a]	0.97 (0.25) [a]	1.97 (0.25) [b]	2.23 (0.25) [b]
Vocabulary	27.33 (0.74) [a]	27.87 (0.74) [a]	34.57 (0.74) [b]	34.33 (0.74) [b]
Digit-Symbol	61.50 (1.83) [a]	57.47 (1.83) [a]	49.90 (1.83) [b]	49.63 (1.83) [b]
D–S memory	7.93 (0.35) [a]	6.76 (0.35) [a]	5.00 (0.35) [b]	4.70 (0.35) [b]

Note. Age differences were significant for all effects, *p*s <.05. Differing superscripts in each row represent significant differences. Standard Errors are reported in parentheses. Vocabulary score was out of 40 (Shipley Institute of Living Scale; Zachary & Shipley, 1986); No. mediations: self-reported number of daily medications; Digit-Symbol (Weschler, 1981):

WAIS-R Digit-Symbol subtest; D–S memory: symbol memory following Digit-Symbol subtest; MP = market-priced condition; OP = overpriced condition.

Table 2

Pearson and Spearman Correlations of Retrieval Use between the Noun-Pair Lookup Task and the Grocery-Pair Lookup Task

	Non-retrievers excluded		Non-retrievers included	
Sample	r [LL, UL]	r_s [LL, UL]	r [LL, UL]	r_s [LL, UL]
Whole sample	.29 [.11, .46]	.25 [.06, .42]	.51 [.36, .63]	.39 [.23, .54]
Young adults	-.01 [-.28, .25]	.14 [-.13, .39]	.37 [.13, .58]	.23 [-.04, .46]
Older adults	.33 [.06, .56]	.31 [.03, .54]	.51 [.29, .68]	.48 [.26, .66]
Market-priced	.19 [-.07, .44]	.21 [-.06, .45]	.37 [.12, .57]	.28 [.02, .49]
Overpriced	.36 [.09, .58]	.29 [.02, .53]	.57 [.37, .73]	.50 [.28, .67]
YA-MP	.003 [-.36, .37]	.23 [-.15, .55]	.003 [-.36, .37]	.23 [-.15, .55]
OA-MP	.25 [-.14, .57]	.25 [-.14, .58]	.41 [.05, .68]	.36 [-.01, .64]
YA-OP	-.05 [-.42, .35]	.04 [-.35, .42]	.47 [.11, .71]	.22 [-.17, .55]
OA-OP	.35 [-.08, .66]	.34 [-.09, .66]	.52 [.19, .75]	.58 [.27, .78]

Note. [LL, UL] = lower and upper limit of the 95% confidence interval around the correlation estimate; r_s = Spearman correlation; r = Pearson correlation. YA-MP = Young adults in the market-priced condition; OA-MP= Older adults in the market-priced condition; YA-OP: Young adults in the overpriced condition; OA-OP: Older adults in the overpriced condition. Non-retrievers were participants who did not endorse the retrieval strategy in the final block of practice in both tasks. Non-retrievers included two participants from the YA-OP condition, two from the OA-MP condition, and six from the OA-OP condition.

Table 3

Post-Task Questionnaire means for the Grocery-Pair and Noun-Pair Lookup Tasks

Question	GPLT	NPLT
1. Estimated total pairs memorized	YA-MP: 8.55 (.49) [a]	YA-MP: 9.93 (0.48) [a]
	OA-MP: 8.03 (.48) [a]	OA-MP: 6.38 (0.48) [b]
	YA-OP: 5.62 (.49) [b]	YA-OP: 9.96 (0.48) [a]
	OA-OP: 5.13 (.48) [b]	OA-OP: 6.75 (0.48) [b]
2. Global confidence	YA-MP: 73.79 (4.86) [a]	YA-MP: 87.86 (4.42) [a]
	OA-MP: 71.07 (4.77) [a]	OA-MP: 53.31 (4.41) [b]
	YA-OP: 48.31 (4.85) [b]	YA-OP: 81.89 (4.48) [a]
	OA-OP: 52.07 (4.77) [b]	OA-OP: 63.68 (4.48) [b]
3. Estimated memory strategy usage	YA-MP: 77.72 (4.71) [a]	YA-MP: 77.93 (4.88) [a]
	OA-MP: 55.53 (4.63) [b]	OA-MP: 34.38 (4.88) [b]
	YA-OP: 48.24 (4.71) [b]	YA-OP: 79.50 (4.97) [a]
	OA-OP: 27.13 (4.63) [c]	OA-OP: 31.00 (4.97) [b]
4. Effort learning pairs	YA-MP: 50.00 (5.19) [a]	YA-MP: 42.06 (5.13) [a]
	OA-MP: 54.73 (5.11) [a]	OA-MP: 46.28 (5.13) [a]
	YA-OP: 66.41 (5.19) [a]	YA-OP: 43.18 (5.22) [a]
	OA-OP: 54.97 (5.11) [a]	OA-OP: 50.89 (5.22) [a]
5. Difficulty of task	YA-MP: 48.97 (4.71) [a]	YA: 23.66 (5.05) [a]
	OA-MP: 38.23 (4.63) [a]	OA-MP: 39.76 (5.05) [b]
	YA-OP: 66.07 (4.71) [b]	YA-OP: 25.89 (5.14) [a]
	OA-OP: 60.47 (4.63) [b]	OA: 35.18 (3.60) [b]
6. Difficulty of Grocery-Pair task compared to NPLT (0 easier; 100, harder)	YA-MP: 68.38 (5.08) [a]	
	OA-MP: 26.50 (5.00) [b]	
	YA-OP: 77.62 (5.08) [a]	
	OA-OP: 59.23 (5.00) [a]	
7. Familiarity with grocery items	YA-MP: 85.72 (4.55) [a]	
	OA-MP: 86.80 (4.47) [a]	

	YA-OP: 79.31 (4.55) ^{a}
	OA-OP: 88.47 (4.47) ^{a}
8. How often do you go grocery shopping?	YA-MP: 67.72 (5.12) ^{a}
	OA-MP: 78.13 (5.03) ^{a}
	YA-OP: 62.86 (5.12) ^{a}
	OA-OP: 78.93 (5.03) ^{a}
9. How well do you think you did overall?	YA-MP: 75.37 (3.75) ^{a}
	OA-MP: 66.50 (3.69) ^{a}
	YA-OP: 62.10 (3.75) ^{a}
	OA-OP: 69.00 (3.69) ^{a}
10. Satisfaction with performance	YA-MP: 77.21 (4.68) ^{a}
	OA-MP: 65.37 (4.61) ^{a}
	YA-OP: 60.97 (4.68) ^{a}
	OA-OP: 61.90 (4.61) ^{a}
11. Could you have done better if you tried harder?	YA-MP: 66.38 (6.27) ^{a}
	OA-MP: 49.17 (6.16) ^{a}
	YA-OP: 56.83 (6.27) ^{a}
	OA-OP: 49.80 (6.16) ^{a}
12. Did you find the experiment fatiguing (0 not fatiguing, 100 fatiguing)?	YA-MP: 70.14 (5.37) ^{a}
	OA-MP: 39.28 (5.27) ^{b}
	YA-OP: 67.00 (5.37) ^{a}
	OA-OP: 41.00 (5.27) ^{b}
13. Did you feel stress or tension during the experiment? (0 no stress, 100 stress)	YA-MP: 24.86 (5.04) ^{a}
	OA-MP: 23.33 (4.95) ^{a}
	YA-OP: 31.41 (5.04) ^{a}
	OA-OP: 22.67 (4.95) ^{a}

Note. YA-MP = Young adults in the market-priced condition; OA-MP = Older adults in the market-priced condition; YA-OP = Young adults in the overpriced condition; OA-OP = Older adults in the overpriced condition. Different superscripts designate significant differences using Tukey's HSD. Only five questions were asked for the Noun-Pair Post-task questionnaire, and no condition differences were evident.

Table 4

Correlations of End-Retrieval Use with Metacognitive Measures

	GPLT		NPLT	
PTQ	YAs	OAs	YAs	OAs
Estimated pairs memorized	.47**	.59**	.52**	.50**
Global confidence	.47**	.57**	.31*	.36*
Estimated use of retrieval strategy	.42**	.71**	.36*	.61**
Effort memorization	-.23	.07	-.06	.39**
Difficulty of task	-.40*	-.25	-.07	.02
GPLT vs NPLT	-.19	-.36*		
Familiarity with groceries	.38*	.17		
Frequency of grocery shopping	.22	.06		
How well do you think you did overall?	.37*	.22		
Satisfaction with performance	.25	.08		
Could you have done better?	.10	-.01		

Was the study fatiguing?	.00	.03		
Did you feel stress or tension?	-.17	.13		
JOLs	.57**	.67**	.49**	.37*
C-R	.54**	.68**	.42**	.55**
d'	.57**	.62**	.79**	.67**
PBMI: Global MSE	.14	.07	.17	-.10
PBMI: Relative standing	.14	.02	.06	-.14
PBMI: Retrospective Change	-.10	-.14	-.13	-.23
PBMI: Prospective Change	-.09	-.10	.00	-.16
PBMI: Control	.15	-.01	.22	-.03
PBMI: Prospective Control	-.07	-.23	.09	-.35*
PBMI: Future Control	-.06	-.14	.05	-.26
PBMI: Specific MSE	.26	-.02	.36*	-.15

Note. Differing alpha thresholds were set between the GPLT and NPLT PTQ correlations due to differing numbers of correlations calculated. For the GPLT correlations, alpha was set to .002 (.05/24); for the NPLT correlations, alpha was set to .00315 (.05/16). Correlations not meeting the corrected alpha thresholds are reported for documentation but were not interpreted due to the exploratory nature of the analyses. GPLT = Grocery-Pair Lookup Task; NPLT = Noun-Pair Lookup

Task; YAs = Young adults; OAs = Older adults; PBMI = Personal Beliefs about Memory Instrument; MSE = Memory Self-Efficacy.

* $p < .05$. ** $p <$ alpha threshold.

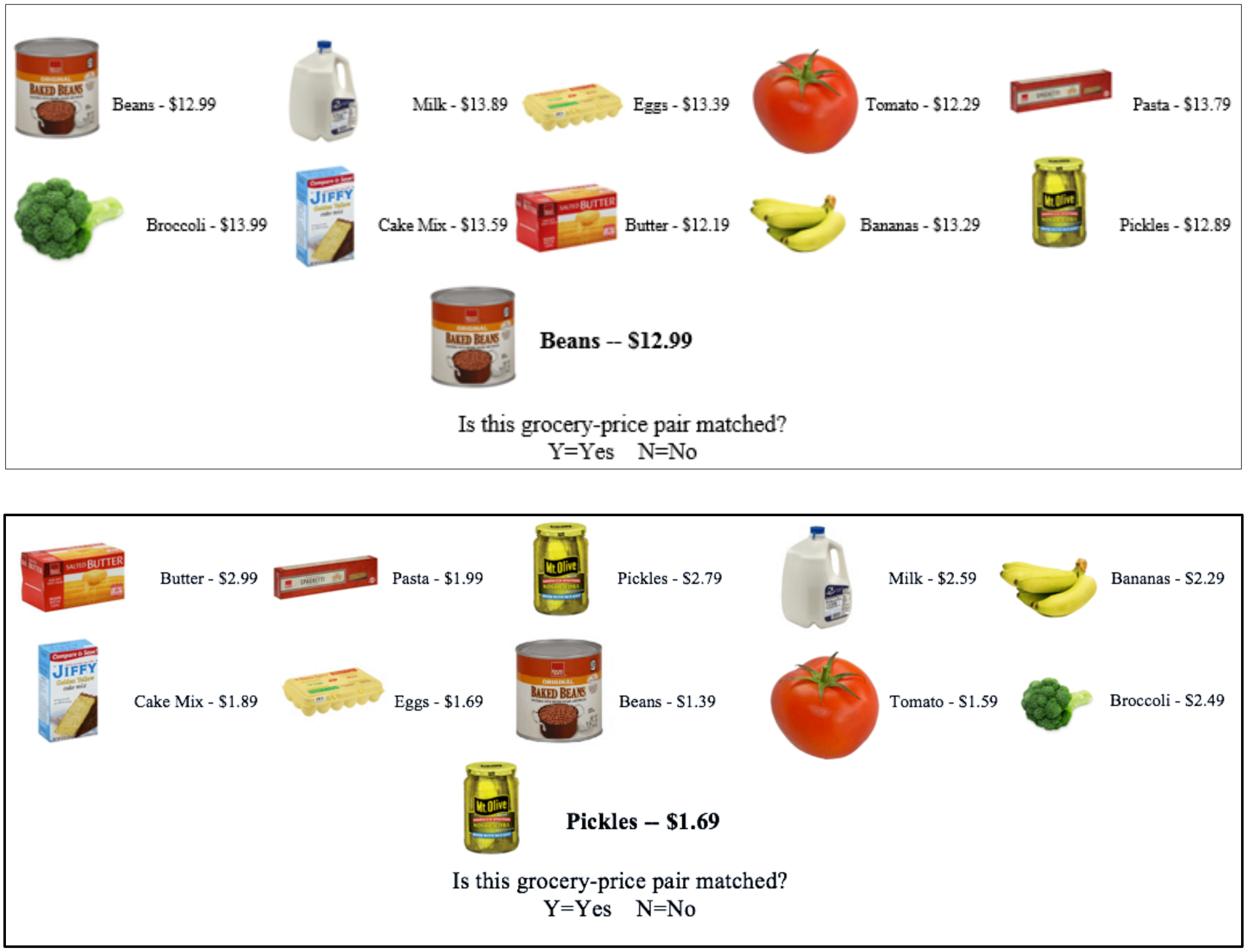

Figure 1. Example of the Grocery-Pair Lookup Task. The top figure represents the overpriced condition on a match trial, and the bottom figure represents the market-priced condition on a mismatch trial.

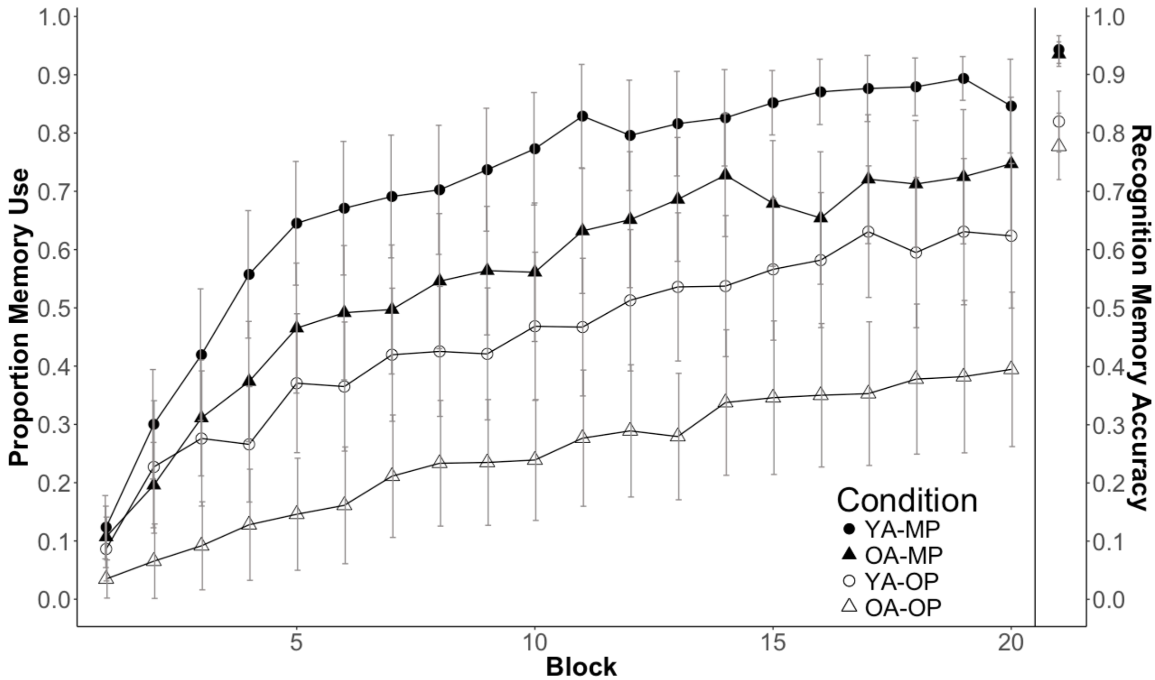

Figure 2. Endorsement of the Retrieval Strategy in the Grocery-Pair Lookup Task. Separate lines and shapes indicating age group and condition, followed by recognition memory accuracy after the vertical line. By the final block, older adults in the market-priced condition were equivalent with young adults in the market-priced condition, and young adults in the overpriced condition. Only condition differences were significant in recognition memory, where the market-priced group performed better than the overpriced group. MP = Market-priced condition; OP = Overpriced condition. Error bars are 95% CI.

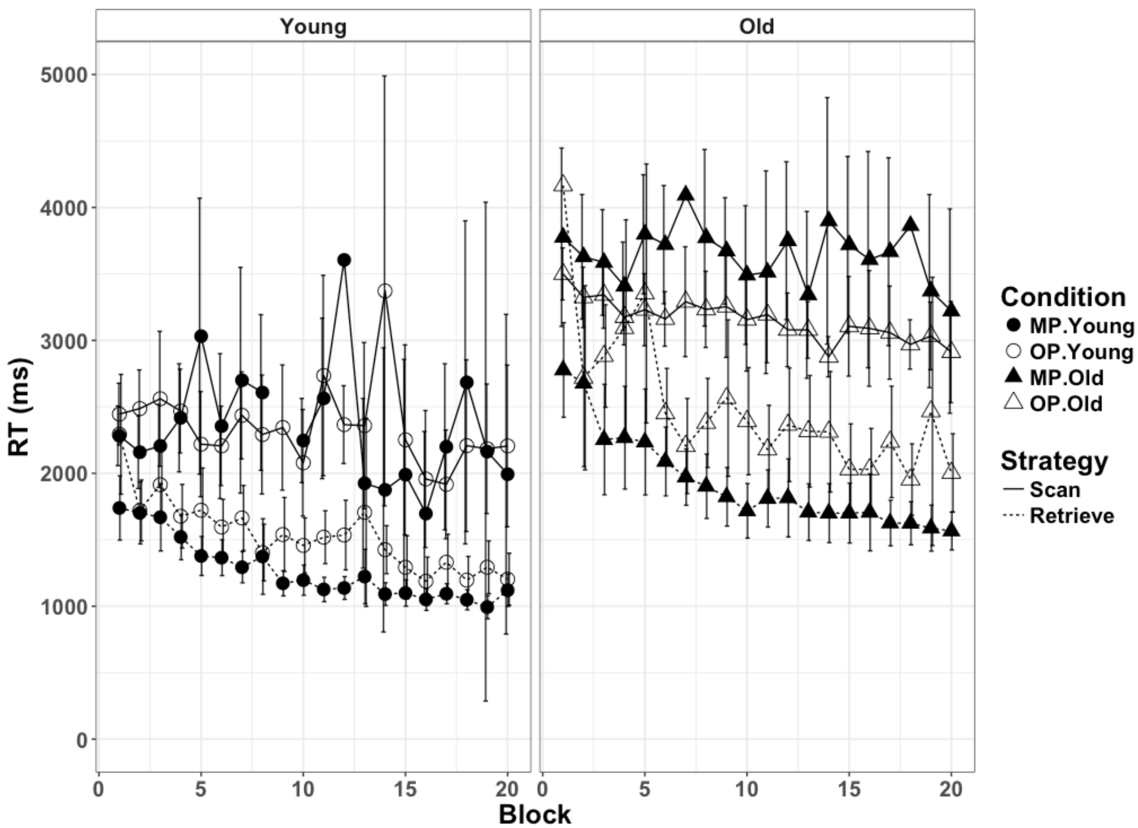

Figure 3. Reaction Times for the Scanning and Retrieval Strategies in the Grocery-Pair Lookup Task. The two panels are separated by age to examine condition differences within age group. Retrieval times improved with practice. Because of changing rates of strategy endorsements, estimates in earlier blocks for scanning are more reliable than later blocks; estimates in later blocks for retrieval are more reliable than earlier blocks. Scanning reaction times remained relatively stable over practice. The range of the figure was restricted for ease of viewing, resulting in missing values and error bars when they occurred outside that range. MP = Market-priced; OP = Overpriced. Error bars represent 95% CI.

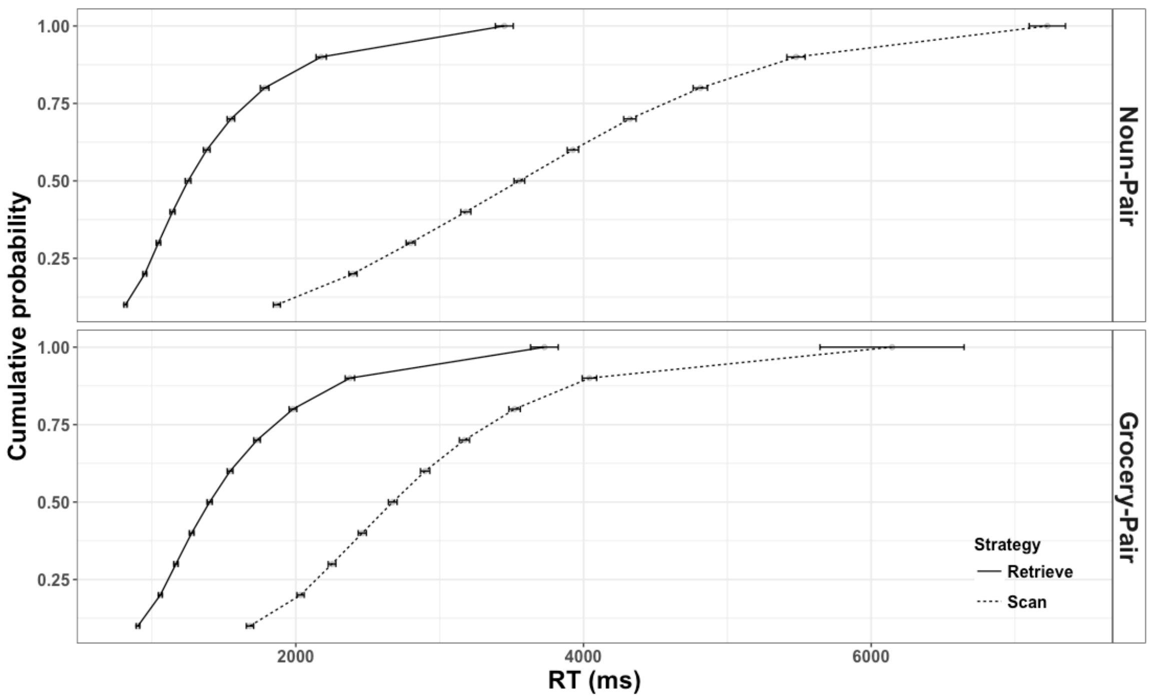

Figure 4. Cumulative Distribution Functions of Reaction Times by Strategy. Rank ordered reaction times for each task were binned into deciles and averaged within participants and then across decile as a means of validating the strategy endorsement options for the NPLT (top) and GPLT (bottom). Lack of crossover between the two functions indicates that they are separate functions, i.e., representing two different cognitive processes, which is clearly demonstrated. Error bars represent 95% CIs.

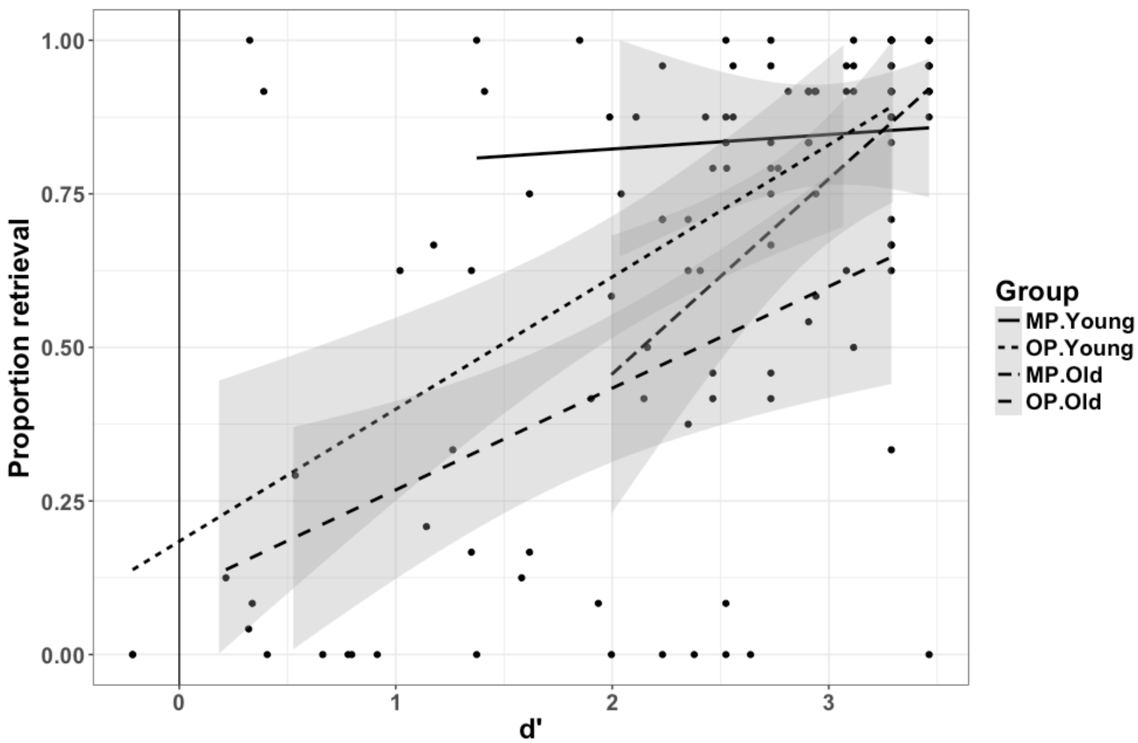

Figure 5. Proportion Retrieval in the Final Block of Practice in the Grocery-Pair Lookup Task Regressed on Age, Condition, and d'. Trend lines are plotted for each group. Chance performance in recognition memory is indicated with the vertical line at 0. Perfect accuracy was corrected to equal ~3.49. MP = market-priced; OP = overpriced. Error bars are 95% CI.

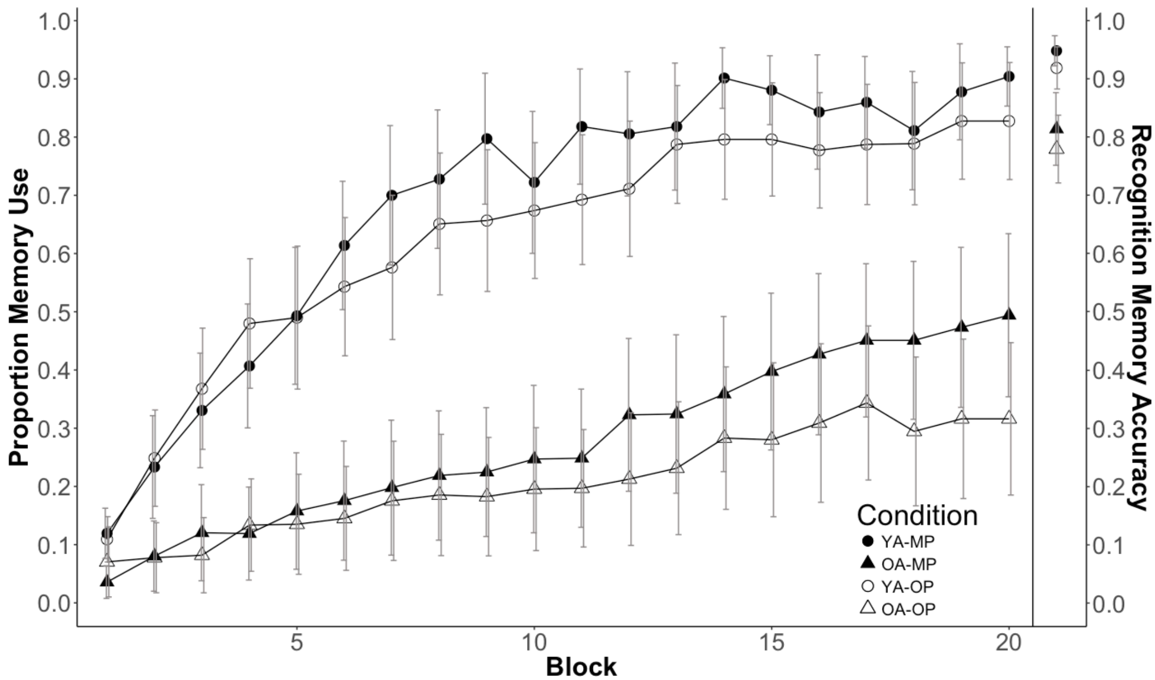

Figure 6. Endorsement of the Retrieval Strategy in the Noun-Pair Lookup Task. The panels are separated by age and condition, followed by recognition memory. Young adults are represented in solid lines, and older adults are represented by dashed lines, respectively. Condition is designated by shape. Post-hoc analyses revealed no significant differences between older adult groups in the final block, $p = .09$. Recognition memory data are plotted following the vertical line. Only the age difference was significant for recognition memory. The condition designation was not meaningful in this task. Error bars are 95% CI. MP = market-priced; OP = overpriced.

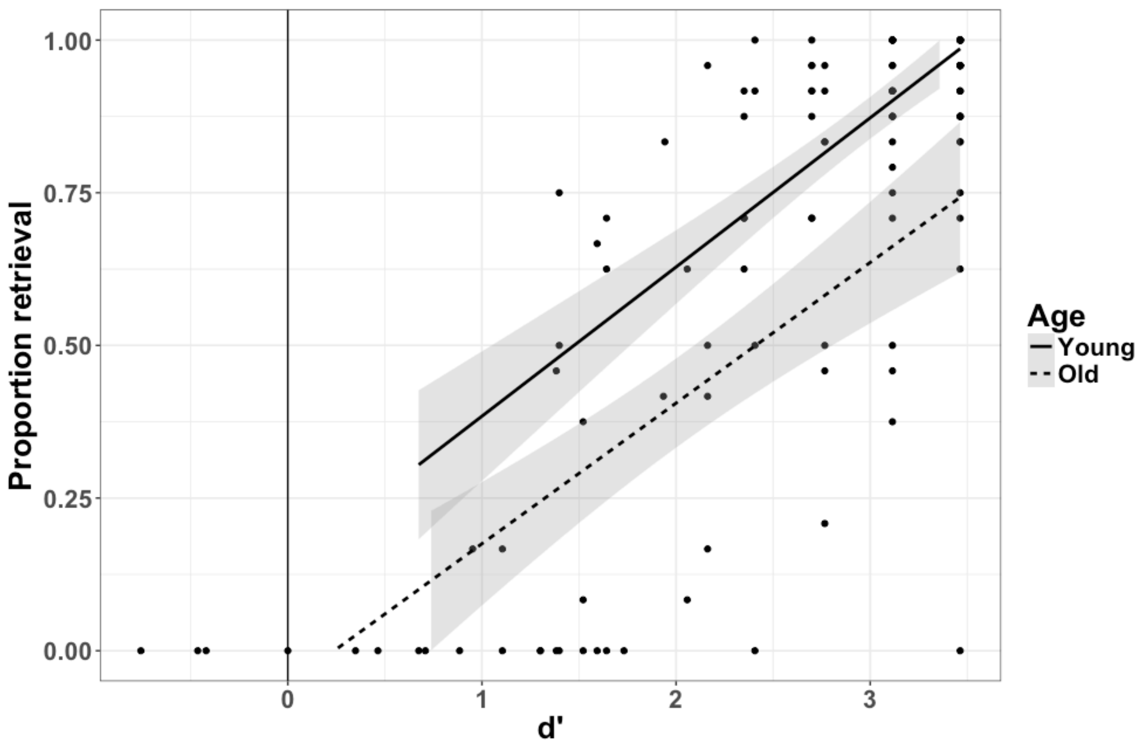

Figure 7. Proportion Retrieval in the Final Block of Noun-Pair Practice Regressed on d'. Trend lines plotted for each age-group. Chance performance in recognition memory is indicated with the vertical line at 0. Error bars are 95% CI.

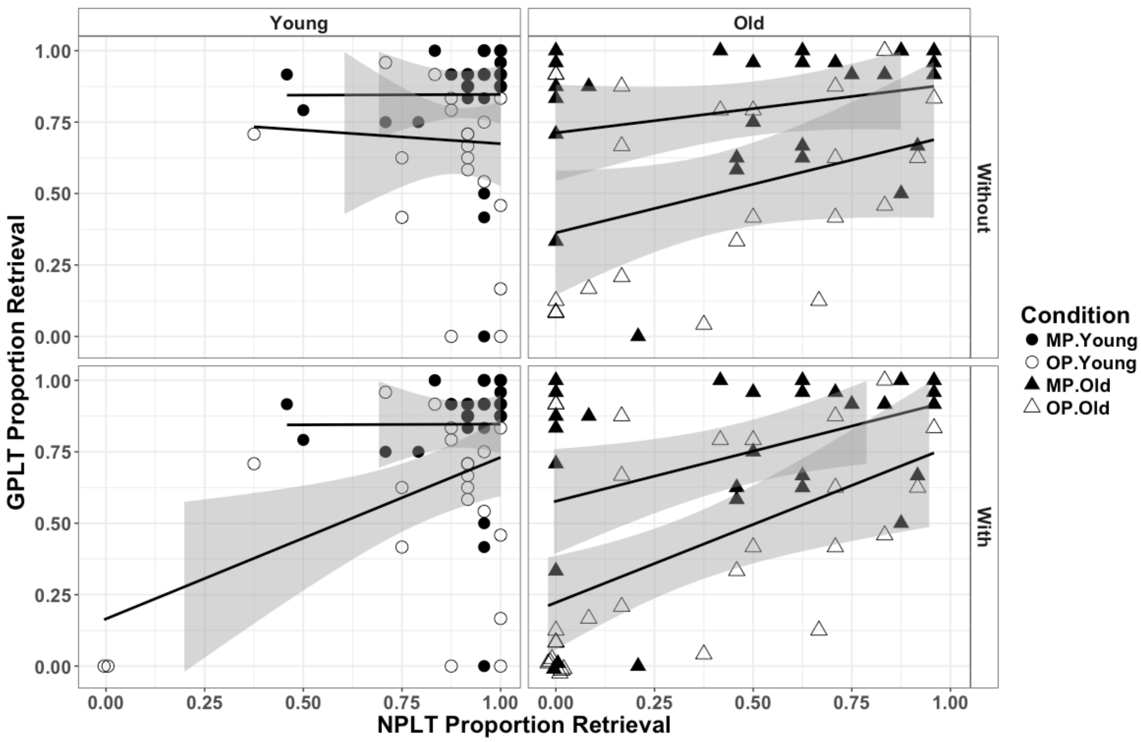

Figure 8. Retrieval Use in the Final Blocks of the Grocery-Pair Lookup Task Regressed on the Noun-Pair Lookup Task. The figure is separated by age, and whether participants that did not endorse retrieval in the final blocks of both tasks were included (With), or excluded (Without) to demonstrate the dramatic effects of their inclusion. No relationship of retrieval use between both tasks is found for young adults, while a modest relationship is suggested for older adults. The points around (0, 0) in the bottom row have been jittered to give a sense of the number of participants that did not retrieve in the final blocks of either task. Shaded regions indicate 95% confidence intervals. MP = Market-priced condition; OP = Overpriced condition.

APPENDIX B

LIST OF POST-TASK QUESTIONNAIRE ITEMS

Table B1

Grocery-Pair Lookup Task

Question Number	Question
1.	How many of the 12 grocery prices do you think you had memorized by the end of the experiment? Please type a number between 0 and 12.
2.	Were you confident of success when using your memory? Please indicate your confidence level with a number between 0 (not confident) and 100 (very confident).
3.	Estimate how often you used the memory-only strategy as a number between 0 (Never) and 100 (Every Time).
4.	Some people learn the prices without even trying, just because they are repeated over and over again. Think about how you learned the prices, if at all. For the prices you did learn, rate how much that learning involved a deliberate effort between 0 (Automatic) to 100 (Effortful).

5.	Rate how difficult you found the Grocery-Pair task. Please rate the difficulty between 0 (Very Easy) and 100 (Very Difficult).
6.	Rate how difficult you found the Grocery-Pair task compared to the Noun-Pair Task. Please rate the difficulty between 0 (Much Easier than the Noun-Pair Task) and 100 (Much More Difficult than the Noun-Pair Task).
7.	How familiar were you with the 12 grocery items that you saw today? Please rate your familiarity between 0 (Not Familiar) and 100 (Very Familiar).
8.	How often do you go grocery shopping? Please answer between 0 (Almost Never) and 100 (Very Often).
9.	How well do you think you did in this study overall? Please rate your confidence between 0 (Not Well) and 100 (Very Well).
10.	Does your performance on the experimental task satisfy you? Please rate your satisfaction between 0 (Not Satisfied) and 100 (Very Satisfied).
11.	Would you be able to do better in this study if you tried harder? Please rate your confidence between 0 (Not Likely) and 100 (Very Likely).
12.	How fatiguing did you find the experimental task overall? Please rate your fatigue between 0 (Not At All) and 100 (Very Fatiguing).
13.	Did you feel stress or tension during this study? Please rate your stress level between 0 (Not At All) and 100 (Very Stressed).

Note. Grocery-Pair Lookup Task followed the Noun-Pair Lookup Task. Questions 9 – 14 regarded the entire study, rather than the Grocery-Pair Lookup Task specifically.

Table B2

Noun-Pair Lookup Task

Question Number	Question
1.	How many of the 12 noun-pairs do you think you had memorized by the end of the experiment? Please type a number between 0 and 12.
2.	Were you confident of success when using your memory? Please indicate your confidence level with a number between 0 (not confident) and 100 (very confident).
3.	Estimate how often you used the memory-only strategy. Please type your answer as a number between 0 (Never) and 100 (Every Time).
4.	Some people learn the pairings without even trying, just because they are repeated over and over again. Think about how you learned the pairs, if at all. For the pairs you did learn, rate how much that learning involved a deliberate effort between 0 (Automatic) to 100 (Effortful).
5.	Rate how difficult you found the computer task. Please rate the difficulty between 0 (Very Easy) and 100 (Very Difficult).

Note. The computer task mentioned in question 5 specifically refers to the Noun-Pair Lookup Task.

APPENDIX C

CORRELATIONS OF PBMI ITEMS WITH POST-TASK QUESTIONNAIRE ITEMS

PBMI items were correlated with post-task questionnaire items in both the Noun-Pair and Grocery-Pair Lookup Tasks, collapsing across age and condition for better power. Correlations are reported in Table C1 for Grocery-Pair items and Table C2 for Noun-Pair items. Grocery-Pair correlations are reported first. Only the first six grocery post-task questionnaire items were used, as they related directly to the task itself, whereas the remaining question-items were concerned with overall task performance. JOLs, cued-recall accuracy, and *d'*, were also included. As such, the nine items were individually correlated with the eight PBMI measures, and my alpha was restricted to .05/72 = .00069, for each correlation. No significant correlations emerged. I shall comment on several theoretically interesting correlations that met traditional significance at .05, nonetheless, in the interest of motivating future hypotheses. Global memory ability, or Global MSE, correlated with estimated number of items memorized, $r = .22$, 95% CI [.04, .39], $p = .02$. Global MSE also correlated with estimated memory strategy use, $r = .26$, 95% CI [.07, .42], $p = .006$. Global MSE also correlated with JOLs, $r = .24$, 95% CI [.06, .41]. It did not correlate with cued recall nor *d'*, *p*s > .08. However, the point estimate for the cued recall correlation was within the average correlation width found by Beaudoin and Desrichard (2011), being $r = .16$, 95% CI [-.02, .34].

Retrospective change correlated with perceived difficulty of the GPLT to the NPLT, $r = .24$, 95% CI [.06, .41], $p = .01$, and with estimated use of the retrieval strategy, $r = .21$, 95% CI [.02, .38], $p = .03$, indicating that the more participants perceived a positive change in their memory abilities over the past 10 years, the easier they found the GPLT in comparison to the NPLT, and the greater the amount of estimated retrieval strategy use. Prospective change also correlated with perceived difficulty of the GPLT to the NPLT, $r = .19$, 95% CI [.001, .36], $p = .049$, and with estimated use of the retrieval strategy, $r = .19$, 95% CI [.003, .36], $p = .046$. Specific MSE correlated with several items: estimated number of memorized items, $r = .21$, 95% CI [.03, .38], $p = .02$; memory strategy confidence, $r = .23$, 95% CI [.05, .40], $p = .01$; and JOLs, $r = .23$, 95% CI [.05, .39].

For correlations with the NPLT, all five post-task questionnaire items, and JOLs, cued-recall, and *d'* were correlated with the PBMI items, for a total of eight measures to correlate with the eight PBMI items, reported in Table C2. I set alpha to be .05/64 = .00078. Three correlations were significant from this series of analyses, along with several that met traditional significance, to be reported later, but not interpreted.

Prospective change significantly correlated with estimated memory strategy use, $r = .34$, 95% CI [.15, .49], $p = .0005$. Prospective change also significantly correlated with JOLs, $r = .33$, 95% CI [.15, .49], $p = .0005$. Those predicting positive changes in their memory in the next 10 years tended to estimate retrieving more, and to be more confident in estimating the likelihood of recalling the noun-pairs. This might be an effect due to age, given that young adults tended to predict more positive change than older adults (*M*

= 45.38, *SE* = 2.04; *M* = 63.91, *SE* = 2.16, for older and young adults' prospective change predictions, closer to 100 representing positive change). Several other correlations met traditional significance when correlated with prospective change: estimated number of memorized pairs, $r = .22$, 95% CI [.03, .39], $p = .02$, memory strategy confidence, $r = .26$, 95% CI [.07, .43], $p = .008$, and cued recall, $r = .21$, 95% CI [.01, .38], $p = .036$.

Retrospective change significantly correlated with estimated memory use in the NPLT, $r = .37$, 95% CI [.19, .53], $p < .001$. Retrospective change also significantly correlated with JOLs, $r = .39$, 95% CI [.21, .54], $p < .001$. Several other items correlated at traditional significance with retrospective change: estimated number of memorized items, $r = .26$, 95% CI [.07, .43], $p = .008$; memory strategy confidence, $r = .29$, 95% CI [.11, .46], $p = .002$; perceived difficulty of the NPLT, $r = -.19$, 95% CI [-.38, -.006], $p = .04$; and cued-recall, $r = .27$, 95% CI [.08, .44], $p = .005$. The correlations with retrospective change might also be driven by age effects, limiting their interpretive and predictive value (*M* = 53.29, *SE* = 1.98; *M* = 78.19, *SE* = 2.08, for older and young adults subjective retrospective change, values at 50 represent no change, values closer to 100 represent positive change).

Other theoretically interesting but not significant correlations included global MSE and memory strategy confidence, $r = .19$, 95% CI [.001, .37], $p = .049$; global MSE and JOLs, $r = .26$, 95% CI [.07, .43], $p = .007$; prospective change and JOLs, $r = .26$, 95% CI [.07, .43], $p = .007$; prospective control over memory (i.e., perceived ability to be able to act now to increase control their memory in the future) and estimated number of memorized items, $r = -.19$, 95% CI [-.38, -.01], $p = .04$; and future control over memory (i.e., perceived ability to control their memory in the future), r = -.20, 95% CI [-.38, -.01]. The pattern of correlations observed across both the Grocery and Noun-Pair Lookup Tasks suggests that MSE tends to correlate moderately with metacognitive measures from these tasks, like estimated number of memorized items, and JOLS. The effects of memory self-efficacy have been proposed to work via motivational factors, not including metacognitive factors, (e.g., Beaudoin & Desrichard, 2011), but the two factors are not the usual targets of investigation (usually MSE is examined in relation to memory). The metacognitive factors that correlated here with the PBMI items here could be mediators in the relationship between metamemory and memory, and could be targeted in future work.

Table C1

Grocery-Pair Lookup Task Post-Task Questionnaire by PBMI Correlations

PBMI	PTQ items								
	Estimated number of pairs memorized	Global confidence in memory strategy	Estimated use of retrieval strategy	Memorization effort	GPLT difficul-ty	GPLT vs NPLT difficulty	JOLs	C-R	*d’*
Global MSE	.22*	.17	.26*	-.04	-.13	-.05	.24*	.16	.05
Relative standing	.17	.12	.16	-.12	-.16	-.07	.18	.08	-.02
Retrospective Change	.04	.01	.21*	.07	.04	.24*	-.01	.01	-.04
Prospective Change	.12	.00	.19*	.17	-.02	.19*	.03	.02	-.01
Control	.17	.14	.14	.04	-.10	-.04	.07	.08	.02
Prospective Control	-.04	-.01	-.06	-.05	.05	.04	-.08	-.12	-.08
Future Control	.00	.03	-.01	.02	.06	.03	-.01	-.04	-.08

Specific MSE	.21*	.23*	.11	.05	-.17	-.03	.23*	.14	.08

Note. C-R = Cued-Recall; PTQ = Post-Task Questionnaire; JOLs = Judgments of Learning; NPLT = Noun-Pair Lookup Task; GPTL = Grocery-Pair Lookup Task; MSE = Memory Self-Efficacy. No correlations met the significance threshold.

* $p < .05$.

Table C2

Noun Pair Post-Task Questionnaire by PBMI Correlations

PBMI	PTQ items							
	Estimated number of pairs memorized	Global confidence in memory strategy	Estimated use of retrieval strategy	Memorization effort	NPLT difficulty	JOLs	C-R	*d'*
Global MSE	.12	.19	.12	.05	-.06	.26*	.10	-.04
Relative standing	-.02	.11	.08	.03	-.08	.14	-.05	-.1
Retrospective Change	.26*	.29*	.37**	-.01	-.20*	.39**	.27*	.23*
Prospective Change	.22*	.26*	.34**	-.02	-.13	.33**	.21*	.11
Control	-.08	.10	.00	-.04	.08	.09	.04	.00
Prospective Control	-.20*	-.09	-.11	-.13	.15	-.09	-.14	-.05
Future Control	-.20*	-.07	-.13	-.04	.16	-.05	-.10	-.03
Specific MSE	.05	.12	.01	.02	-.09	.14	.00	-.01

Note. C-R = Cued-Recall; PTQ = Post-Task Questionnaire; JOLs = Judgments of Learning; NPLT = Noun-Pair Lookup Task; MSE = Memory Self-Efficacy.

* $p < .05$. ** $p < .0007$, the significance threshold